

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:04:14 ; Search time 20.2998 Seconds
(without alignments)
483.460 Million cell updates/sec

Title: US-09-758-308-2
Perfect score: 536
Sequence: 1 VPHNIEEVALSNTGIEIPFY.....LMTGTGDFSDVIDNCTCVT 102

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	525.5	98.0	3010	1 GNVVCJ	genome polyprotein
2	514.5	96.0	3010	1 A45573	genome polyprotein
3	507.5	94.7	3010	1 GNVVTC	genome polyprotein
4	504.5	94.1	216	2 S21337	genome polyprotein
5	498.5	93.0	3010	1 S18030	genome polyprotein
6	497.5	92.8	3011	1 GNVVC3	genome polyprotein
7	493.5	92.1	3010	1 GNVVTV	genome polyprotein
8	492.5	91.9	386	2 S68016	ATPase/RNA helicase
9	492.5	91.9	3011	1 S40770	genome polyprotein
10	489.5	91.3	3011	1 GNVVCH	genome polyprotein
11	457.5	85.4	3033	1 JQ1303	genome polyprotein
12	446.5	83.3	876	2 PC3219	polypeptide - hepa
13	442.5	82.6	3033	1 GNVVJ8	genome polyprotein
14	381.5	71.2	3014	1 JC5620	genome polyprotein
15	296.5	55.3	102	2 PC5028	genome polyprotein
16	260.5	48.6	2970	2 T08839	polyprotein - maur
17	253.5	47.3	3005	2 T08841	polyprotein - dour
18	109.5	20.4	3898	1 A44217	genome polyprotein
19	109.5	20.4	3898	1 GNVVTV	genome polyprotein
20	104.5	19.5	3898	1 GNVVHC	genome polyprotein
21	104.5	19.5	3898	1 GNVVHB	genome polyprotein
22	104.5	19.5	3898	2 S57437	genome polyprotein
23	104	19.4	3898	2 S58295	polyprotein - hog
24	101	18.8	855	2 A71223	hypothetical prote
25	101	18.8	855	2 B75191	probable ATP-depen
26	87.5	16.3	610	1 BVECRQ	DNA helicase recQ
27	87.5	16.3	611	2 F86069	ATP-dependent DNA
28	87.5	16.3	611	2 H91222	ATP-dependent DNA
29	87.5	16.3	766	2 H81794	ATP-dependent DNA

RESULT 1
GNVVCJ

genome polyprotein - hepatitis C virus (strain J)
N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A39253; PS0086
R:Kato, N., Hijikata, M., Ootsuyama, Y., Nakagawa, M., Ohkoshi, S., Sugimura, T., Shimotohno, K., Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990.
A:Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients
A:Reference number: A39253; MUID:91088550; PMID:2175903
A:Accession: A39253
A:Molecule type: genomic RNA
A:Residues: 1-3010 <KAT>
A:Cross-references: UNIPROT:P26662; GB:D90208; NID:G221610; PID:BAAI4233.1; PID:G221611
R:Kato, N.; Ohkoshi, S.; Shimotohno, K.
Proc. Jpn. Acad. 65B, 219-223, 1989.
A:Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence variations
A:Reference number: PS0085
A:Accession: PS0086
A:Molecule type: genomic RNA
A:Residues: 2650-2707 <KAT>

A:Experimental source: Japanese isolate
C:Comment: The cleavage sites of this polyprotein have not been determined.
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine protease
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <MEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus nonstructural protein NS3 #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4a>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4b>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,

Query Match 98.0% ; Score 525.5 ; DB 1 ; Length 3010;
Best Local Similarity 99.0% ; Pred. No. 9.7e-46;
Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 VPHNIEEVALSNTGIEIPFYKAIPIEAIKGRHLIFCHSKKKDELAALGLGLNAVA 60
Db 1357 VPHNIEEVALSNTGIEIPFYKAIPIEAIKGRHLIFCHSKKKDELAALGLGLNAVA 1416
QY 61 YYRGLDVSVIPISGD-VVVATDALTMTGTFGDFSDVIDNCTCVT 102
Db 1417 YYRGLDVSVIPISGDVVVATDALTMTGTFGDFSDVIDNCTCVT 1459

ATP-dependent DNA
ATP-dependent DNA
hypothetical prote
hypothetical prote
ATP-dependent DNA
probable ATP-depen
DNA helicase XF138
hypothetical prote
DNA helicase TPS1
ATP-dependent DNA
probable helicase
ATP-dependent RNA
ATP-dependent RNA
ATP-dependent RNA
hypothetical prote

RESULT 2

A45573
genome polyprotein - hepatitis C virus (strain JM)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome from a single Japanese carrier: envelope protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: A45573

R.Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata, A. Virus Res. 23, 39-53, 1992

A:Title: Molecular Cloning of hepatitis C virus genome from a single Japanese carrier: envelope protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

A:Reference number: A45573; MUID:92295714; PMID:1318627

A:Accession: A45573

A:Status: preliminary

A:Accession: A45573

A:Molecule type: DNA

A:Residues: 1-3010 <TAK>

A:Cross-references: UNIPROT:Q00269; GB:D011168; GB:D011171; NID:g221612; PIDN:BAA01943.1;

A:Experimental source: HCV-JT

A:Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin

F:2-115/Product: capsid protein C #status predicted <CPC>

F:116-191/Product: envelope protein M #status predicted <EPM>

F:192-383/Product: major envelope protein E #status predicted <ME>

F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>

F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>

F:1007-1615/Product: hepatitis C virus genome polyprotein

F:1230-1337/Region: nucleotide-binding motif A (P-loop)

F:1312-1317/Region: nucleotide-binding motif B

F:1316-1319/Region: DEXH motif

F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>

F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>

F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match 96.0%; Score 514.5; DB 1; Length 3010;

Best Local Similarity 95.1%; Pred. No. 1.3e-44;

Matches 98; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 96.0%; Score 514.5; DB 1; Length 3010;

Best Local Similarity 95.1%; Pred. No. 1.3e-44;

Matches 98; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 96.0%; Score 514.5; DB 1; Length 3010;

Best Local Similarity 95.1%; Pred. No. 1.3e-44;

Matches 98; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 96.0%; Score 514.5; DB 1; Length 3010;

Best Local Similarity 95.1%; Pred. No. 1.3e-44;

Matches 98; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 96.0%; Score 514.5; DB 1; Length 3010;

Best Local Similarity 95.1%; Pred. No. 1.3e-44;

Matches 98; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VATDALMTGFTGDFDSDVIDNCVT 1459

F:1230-1237/Region: nucleotide-binding motif A (P-loop)

F:1312-1317/Region: nucleotide-binding motif B

F:1316-1319/Region: DEXH motif

F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>

F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>

F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

F:196, 209, 234, 250, 305, 325, 417, 423, 430, 448, 532, 540, 556, 576, 623, 645, 1213, 1255, 2041, 2077, 22

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

F:1230-1237/Region: nucleotide-binding motif A (P-loop)

F:1312-1317/Region: nucleotide-binding motif B

F:1316-1319/Region: DEXH motif

F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>

F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>

F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

F:196, 209, 234, 250, 305, 325, 417, 423, 430, 448, 532, 540, 556, 576, 623, 645, 1213, 1255, 2041, 2077, 22

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

QY 61 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 102

DB 1417 YRGLDVSIVPTSGD-VVATDALMTGFTGDFDSDVIDNCVT 1459

Query Match 94.7%; Score 507.5; DB 1; Length 3010;

Best Local Similarity 94.2%; Pred. No. 7.1e-44;

Matches 97; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60

DB 1357 VPHNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 1416

A/Accession: S33570
A/Molecule type: genomic RNA
A/Residues: 1-547, 'T', 549-621, 'V', 623-624, 'S', 626-652, 'DL', 655-761, 'T', 763-782 <HOW>
A/Cross-references: EMBL:X61591
A/Note: this sequence is inconsistent with the nucleotide translation
as Trp, and TTC for residue 771 as Ser
A/Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBI:P:121748)
C/Superfamily: hepatitis C virus genome polyprotein
C/Keywords: ATP; glycoprotein; hydroxylase; nucleotide binding; P-loop; polyprotein; serin
F:115-Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <ME>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polyprotein
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate (As
Query Match 93.0%; Score 498.5; DB 1; Length 3010;
Best Local Similarity 92.2%; Pred. No. 6.1e-43;
Matches 95; Conservative 4; Mismatches 3; Indels 1; Gaps 1;
QY 1 VPHNIEEVALNTGTEIPFYKAIPIEAIKGRHLIFCHSKKKCDLAALKTGLGNVA 60
DB 1357 VPHNIEEVALNTGTEIPFYKAIPIEAIKGRHLIFCHSKKKCDLAALKTGLGNVA 1416
QY 61 YYRGLDVSIVPTSGD-VVATDALTMTGFTGDPDSVIDCNCVT 102
DB 1417 YYRGLDVSIVPTSGD-VVATDALTMTGFTGDPDSVIDCNCVT 1459
RESULT 6
GNWVC3
genome polyprotein - hepatitis C virus (strain HCV-1)
N/Contains: capsid protein C; envelope protein M; hepatitis C virus genome polyprotein NS5
C/Species: hepatitis C virus
C/Date: 30-Sep-1992 #sequence revision 30-Sep-1992 #text_change 09-Jul-2004
C/Accession: A39166; PQ0403; PQ0404
R/Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Coi
Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
A/Title: Genetic organization and diversity of the hepatitis C virus.
A/Reference number: A39166; MUID:91172826; PMID:1848704
A/Accession: A39166
A/Molecule type: mRNA
A/Residues: 1-3011 <CHO>
A/Cross-references: UNIPROT:P26664; GB:M62321; PID:G329873; PID:AAA45676.1; PID:G329874
R/Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.L
J. Gen. Virol. 73, 1131-1141, 1992
A/Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e
A/Reference number: PQ0393; MUID:92268871; PMID:1316939
A/Accession: PQ0403
A/Molecule type: genomic RNA
A/Residues: 1577-1633 <CHA>
A/Cross-references: DDBJ:D10128
A/Experimental source: isolates E-b16
A/Accession: PQ0404
A/Status: preliminary
A/Molecule type: genomic RNA
A/Residues: 1577-1633 <CH2>
A/Experimental source: isolates E-b17
C/Superfamily: hepatitis C virus genome polyprotein
C/Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydroxylase; nonstructura
F:115-Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <ME>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>

F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polyprotein NS5
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2077,222
Query Match 92.8%; Score 497.5; DB 1; Length 3011;
Best Local Similarity 92.2%; Pred. No. 7.7e-43;
Matches 95; Conservative 3; Mismatches 4; Indels 1; Gaps 1;
QY 1 VPHNIEEVALNTGTEIPFYKAIPIEAIKGRHLIFCHSKKKCDLAALKTGLGNVA 60
DB 1357 VPHNIEEVALNTGTEIPFYKAIPIEAIKGRHLIFCHSKKKCDLAALKTGLGNVA 1416
QY 61 YYRGLDVSIVPTSGD-VVATDALTMTGFTGDPDSVIDCNCVT 102
DB 1417 YYRGLDVSIVPTSGD-VVATDALTMTGFTGDPDSVIDCNCVT 1459
RESULT 7
GNWVTW
genome polyprotein - hepatitis C virus (strain Taiwan)
N/Contains: capsid protein C; envelope protein M; hepatitis C virus genome polyprotein NS5
C/Species: hepatitis C virus
A/Note: host Homo sapiens (man)
C/Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 09-Jul-2004
C/Accession: A40244
R/Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
A/Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the
A/Reference number: A40244; MUID:92230206; PMID:1314449
A/Accession: A40244
A/Molecule type: genomic RNA
A/Residues: 1-3010 <CHE>
A/Cross-references: UNIPROT:P29846; GB:M84754
C/Superfamily: hepatitis C virus genome polyprotein
C/Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydroxylase; nonstructura
F:115-Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <ME>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polyprotein NS5
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,222
Query Match 92.1%; Score 493.5; DB 1; Length 3010;
Best Local Similarity 91.3%; Pred. No. 2e-42;
Matches 94; Conservative 5; Mismatches 3; Indels 1; Gaps 1;
QY 1 VPHNIEEVALNTGTEIPFYKAIPIEAIKGRHLIFCHSKKKCDLAALKTGLGNVA 60
DB 1357 VPHNIEEVALNTGTEIPFYKAIPIEAIKGRHLIFCHSKKKCDLAALKTGLGNVA 1416
QY 61 YYRGLDVSIVPTSGD-VVATDALTMTGFTGDPDSVIDCNCVT 102
DB 1417 YYRGLDVSIVPTSGD-VVATDALTMTGFTGDPDSVIDCNCVT 1459
RESULT 8
S68016
ATPase/RNA helicase - hepatitis C virus (fragment)
C/Species: hepatitis C virus

C>Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 09-Jul-2004
C:Accession: S68016
R.Jin, L.; Peterson, D.L.
Arch. Biochem. Biophys. 323, 47-53, 1995
A>Title: Expression, isolation, and characterization of the hepatitis C virus ATPase/RNA
A:Reference number: S68016; MUID:96019946; PMID:7487072
A:Accession: S68016
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-386 <JIN>
A:Cross-references: UNIPROT:Q04045
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; nonstructural protein; nucleotide binding; P-loop; polyprotein
F:24-31/Region: nucleotide-binding motif A (P-loop)
F:86-91/Region: nucleotide-binding motif B
F:90-93/Region: DEXH motif

Query Match 91.9%; Score 492.5; DB 2; Length 386;
Best Local Similarity 91.3%; Pred. No. 2.8e-43;
Matches 94; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKCKDELAALKTGLGNAVA 60
DB 131 VPHNIEEVALSTTGEIPFYGKAIPLEAIKGRHLIFCHSKKCKDELAALKTGLGNAVA 190

QY 61 YYRGLDVSVIPSGD-VVATDALMTGFTGDFSDVIDNCVT 102
DB 191 YYRGLDVSVIPSGD-VVATDALMTGFTGDFSDVIDNCVT 233

RESULT 9
S40770
genome polyprotein - hepatitis C virus
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (strain H)
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Dates: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: S40770; PC1285
R:Okamoto, H.
submitted to the EMBL Data Library, March 1992
A:Reference number: S40770
A:Accession: S40770
A:Molecule type: genomic RNA
A:Residues: 1-3011 <OKA>
A:Cross-references: UNIPROT:Q03463; EMBL:D10749; NID:g221586; PIDN:BAA01582.1; PID:g2215
R:Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda,
Jpn. J. Exp. Med. 60, 167-177, 1990
A>Title: The 5'-terminal sequence of the hepatitis C virus genome.
A:Reference number: PC1284; MUID:91013116; PMID:2170712
A:Accession: PC1285
A:Molecule type: genomic RNA
A:Residues: 1-513 <OK2>
A:Cross-references: GB:D00831; NID:g221511; PIDN:BAA00705.1; PID:g221512
A:Experimental source: isolate HC-J1
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <MEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus (strain H) #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match 91.9%; Score 492.5; DB 1; Length 3011;
Best Local Similarity 92.2%; Pred. No. 2.5e-42;
Matches 95; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKCKDELAALKTGLGNAVA 60
DB 1357 VPHNIEEVALSTTGEIPFYGKAIPLEAIKGRHLIFCHSKKCKDELAALKTGLGNAVA 1416

QY 61 YYRGLDVSVIPSGD-VVATDALMTGFTGDFSDVIDNCVT 102
DB 1417 YYRGLDVSVIPSGD-VVATDALMTGFTGDFSDVIDNCVT 1459

RESULT 10
GNMVCH
genome polyprotein - hepatitis C virus (strain H)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (strain H)
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
A>Note: host Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: A36814; A41546
R:Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
submitted to GenBank, July 1992
A:Description: Genomic structure of the human prototype strain H of hepatitis C virus:
A:Reference number: A36814
A:Accession: A36814
A:Molecule type: genomic RNA
A:Residues: 1-3011 <INC>
A:Cross-references: UNIPROT:P27958; GB:M67463; NID:g329737; PIDN:AAA45534.1; PID:g329737
R:Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A>Title: Genomic structure of the human prototype strain H of hepatitis C virus: compari
A:Reference number: A41546; MUID:92052256; PMID:1658800
A:Contents: annotation
A>Note: neither amino acid nor nucleotide sequence is given
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
F:1-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <MEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus (strain H) #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2240,23

Query Match 91.3%; Score 489.5; DB 1; Length 3011;
Best Local Similarity 91.3%; Pred. No. 5.2e-42;
Matches 94; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKCKDELAALKTGLGNAVA 60
DB 1357 VSHPNIEEVALSTTGEIPFYGKAIPLEIVIKGRHLIFCHSKKCKDELAALKTGLGNAVA 1416

QY 61 YYRGLDVSVIPSGD-VVATDALMTGFTGDFSDVIDNCVT 102
DB 1417 YYRGLDVSVIPSGD-VVATDALMTGFTGDFSDVIDNCVT 1459

RESULT 11
JQ1303
genome polyprotein - hepatitis C virus (isolate HC-J6)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (strain H)
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Dates: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: JQ1303
R:Okamoto, H.; Okada, S.; Sugiyama, Y.; Kurai, K.; Iizuka, H.; Machida, A.; Miyakawa, Y.
J. Gen. Virol. 72, 2697-2704, 1991
A>Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a hum
A:Reference number: JQ1303; MUID:92044440; PMID:1658196


```
A:Accession: JQ1303
A:Molecule type: genomic RNA
A:Residues: 1-3033 <OKA>
A:Cross-references: UNIPROT:P26660; GB:D00944; NID:9221650; PIDN:BAA00792.1; PID:9221651
A:Experimental source: isolate HC-J6 from a Japanese individual
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; glycoprotein; hydrolase; P-loop; polyprotein; serine proteinase; transmembrane
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <MEB>
F:390-733/Product: nonstructural protein NS1 #status predicted <NS1>
F:734-1010/Product: nonstructural protein NS2 #status predicted <NS2>
F:1011-1619/Product: hepatitis C virus genome polyprotein NS4 #status predicted <NS4>
F:1316-1321/Region: nucleotide-binding motif B
F:1320-1323/Region: DEXH motif
F:1620-1866/Product: nonstructural protein NS4a #status predicted <NS4a>
F:1867-2017/Product: nonstructural protein NS4b #status predicted <NS4b>
F:2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038,2888
Query Match 85.4%; Score 457.5; DB 1; Length 3033;
Best Local Similarity 86.3%; Pred. No. 1.1e-38; Mismatches 10; Indels 1; Gaps 1;
Matches 88; Conservative 3;

QY 2 PNPTEEVALSNTGPIPFYKGAIPTEAKGGHRLIFCHSKKCKDELAALRGMLNAVAY 61
DB 1362 PNPTEEVALGEGEIPFVGRAPLSYIKGGHRLIFCHSKKCKDELAALRGMLNAVAY 1421

QY 62 YRGLDVSVIPTSGD-VVWATDALMTGTFGDFSDVDCNVCVT 102
DB 1422 YRGLDVSVIPTQGVVVWATDALMTGTFGDFSDVDCNVAVS 1463

RESULT 12
PC2219
polyprotein - hepatitis C virus (type 5a) (fragments)
N:Contains: core protein; E1 (carboxyl end); E2/NS1 (amino end); NS3 protein; NS4A protein
C:Species: hepatitis C virus
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: PC2219
R:Stuyver, L.; Arnhen, W.V.; Wyseur, A.; Maertens, G.
Biochem. Biophys. Res. Commun. 202, 1308-1314, 1994
A:Title: Cloning and phylogenetic analysis of the core, E2, and NS3/NS4 regions of the H
A:Reference number: PC2219; MUID:94338342; PMID:7520237
A:Accession: PC2219
A:Molecule type: mRNA
A:Residues: 1-876 <STU>
A:Cross-references: UNIPROT:081242; GB:L29577; GB:L29578; GB:L29579
A:Experimental source: serum
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: glycoprotein
F:1-191/Product: core #status predicted <COE>
F:68-78/Region: variable
F:192-247/Product: E1 (carboxyl end) #status predicted <ERE>
F:248-411/Product: E2/NS1 (amino end) #status predicted <ENR>
F:248-338/Region: E2
F:339-411/Region: NS1 (amino end)
F:412-783/Product: NS3 #status predicted <NSR>
F:784-837/Product: NS4A #status predicted <NSA>
F:838-876/Product: NS4B #status predicted <NSB>
F:281,287,294,312,340/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 83.3%; Score 446.5; DB 2; Length 876;
Best Local Similarity 82.4%; Pred. No. 3.9e-38; Mismatches 8; Indels 1; Gaps 1;
Matches 84; Conservative 9;

QY 2 PNPTEEVALSNTGPIPFYKGAIPTEAKGGHRLIFCHSKKCKDELAALRGMLNAVAY 61
DB 484 PNPTEEVALPOGEVFFVGRAPLSYIKGGHRLIFCHSKKCKDELAALRGMLNAVAY 543

QY 62 YRGLDVSVIPTSGD-VVWATDALMTGTFGDFSDVDCNVCVT 102
DB 544 YRGLDVAVPTAGDVVVCSTDALMTGTFGDFSDVDCNSAVT 585
```

RESULT 13

GNWJ8

```
genome polyprotein - hepatitis C virus (strain HC-J8)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome polyprotein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: A40250; PQ0397; PQ0559
R:Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fukuda, S.;
Virology 189, 331-341, 1992
A:Title: Full-length sequence of a hepatitis C virus genome having poor homology to rep
A:Reference number: A40250; MUID:92230232; PMID:1314459
A:Accession: A40250
A:Molecule type: genomic RNA
A:Residues: 1-3033 <OKA>
A:Cross-references: UNIPROT:P26661; GB:D10988; GB:D01221; NID:9221608; PIDN:BAA01761.1;
R:Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.;
J. Gen. Virol. 73, 1131-1141, 1992
A:Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to
A:Reference number: PQ0393; MUID:92268871; PMID:1316939
A:Accession: PQ0397
A:Molecule type: genomic RNA
A:Residues: 2678-2754 <CHA>
A:Cross-references: DDBJ:D10134
A:Experimental source: isolate E-b12
R:Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno
Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A:Title: Distribution of plural HCV types in Japan.
A:Reference number: PQ0554; MUID:92068204; PMID:1720309
A:Accession: PQ0559
A:Molecule type: mRNA
A:Residues: 2678-2729 <KAT>
A:Cross-references: GB:D10562; GB:D90518; NID:9221523; PIDN:BAA01418.1; PID:9221524
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural
F:1-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <MEB>
F:390-733/Product: nonstructural protein NS1 #status predicted <NS1>
F:734-1010/Product: nonstructural protein NS2 #status predicted <NS2>
F:1011-1619/Product: hepatitis C virus genome polyprotein NS4 #status predicted <NS4>
F:1234-1241/Region: nucleotide-binding motif A (P-loop)
F:1316-1321/Region: DEXH motif
F:1320-1323/Region: DEXH motif
F:1620-1866/Product: nonstructural protein NS4a #status predicted <NS4a>
F:1867-2017/Product: nonstructural protein NS4b #status predicted <NS4b>
F:2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,233,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038,2
Query Match 82.6%; Score 442.5; DB 1; Length 3033;
Best Local Similarity 83.3%; Pred. No. 3.8e-37; Mismatches 6; Indels 1; Gaps 1;
Matches 85; Conservative 6;

QY 2 PNPTEEVALSNTGPIPFYKGAIPTEAKGGHRLIFCHSKKCKDELAALRGMLNAVAY 61
DB 1362 PNPTEEVALGEGEIPFVGRAPLSYIKGGHRLIFCHSKKCKDELAALRGMLNAVAY 1421

QY 62 YRGLDVSVIPTSGD-VVWATDALMTGTFGDFSDVDCNVCVT 102
DB 1422 YRGLDVSVIPTQGVVVWATDALMTGTFGDFSDVDCNVAVS 1463

RESULT 14
JCS620
genome polyprotein - hepatitis C virus (isolate EUH1480)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome polyprotein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
A:Accession: JCS620
R:Chamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
```

Biochem. Biophys. Res. Commun. 236, 44-49, 1997
 A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predominant
 A;Reference number: JC5620; MUID:97366593; PMID:9223423
 A;Accession: JC5620

A;Molecule type: mRNA
 A;Residues: 1-3014 <CHA>
 A;Cross-references: UNIPROT:O39928; GB:Y13184
 A;Experimental source: genotype 5a, which predominates in South Africa
 A;Note: the translation of the nucleotide sequence is not complete in this paper
 C;Superfamily: hepatitis C virus genome polyprotein
 C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
 F;2-115/Product: capsid protein C #status predicted <CPC>
 F;116-191/Product: envelope protein M #status predicted <EPM>
 F;192-389/Product: major envelope protein E #status predicted <MEE>
 F;384-408/Region: hypervariable #status predicted
 F;390-730/Product: nonstructural protein NS1 #status predicted <NS1>
 F;731-1007/Product: nonstructural protein NS2 #status predicted <NS2>
 F;1008-1616/Product: hepatitis virus #status predicted <NS3>
 F;1231-1238/Region: nucleotide-binding motif A (P-loop)
 F;1313-1318/Region: nucleotide-binding motif B
 F;1317-1320/Region: DEXH motif
 F;1617-1863/Product: nonstructural protein NS4a #status predicted <N4A>
 F;1864-2014/Product: nonstructural protein NS4b #status predicted <N4B>
 F;2015-3014/Product: nonstructural protein NS5 #status predicted <NS5>
 F;2210-2249/Region: interferon sensitivity determining #status predicted

Query Match 71.2%; Score 381.5; DB 1; Length 3014;
 Best Local Similarity 73.5%; Pred. No. 7.8e-31;
 Matches 75; Conservative 9; Mismatches 17; Indels 1; Gaps 1;

QY 2 PHNTEEVNLTGTEIPYGVKAIPTEAKGSHLIFCHSKKCDLAAKLTGLGNVAY 61
 DB 1359 PHNTEEVNLTGTEIPYGVKAIPTEAKGSHLIFCHSKKCDLAAKLTGLGNVAY 1418

QY 62 YRGLDVSIVPTSGDVVW-ATDALMTGFTGDFSDVIDCNTCVT 102
 DB 1419 YRGLDVSIVPTSGDVVW-ATDALMTGFTGDFSDVIDCNTCVT 1460

RESULT 15

PC6028
 genome polyprotein - hepatitis GB virus C (fragment)
 C;Species: hepatitis GB virus C
 C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 09-Jul-2004
 C;Accession: PC6028
 R;Wang, X.T.; Zhuang, H.; Li, H.; Fan, J.; Qi, Z.; Liu, G.
 Chinese J. Microbiol. Immunol. 16, 263-266, 1996
 A;Title: Detection of GBV-C infection and sequencing of partial gene of a Chinese strain
 A;Reference number: PC6028
 A;Accession: PC6028
 A;Molecule type: mRNA
 A;Residues: 1-102 <WAN>
 A;Cross-references: UNIPROT:Q69899; UNIPROT:Q69423; UNIPROT:O09803; UNIPROT:O36180; UNIP
 A;Note: the authors translated the codon GGT for residue 32 as Thr
 C;Superfamily: hepatitis C virus genome polyprotein
 C;Keywords: polyprotein

Query Match 55.3%; Score 296.5; DB 2; Length 102;
 Best Local Similarity 56.6%; Pred. No. 1.3e-23;
 Matches 56; Conservative 15; Mismatches 27; Indels 1; Gaps 1;

QY 3 HPIIEEVALNTGTEIPYGVKAIPTEAKGSHLIFCHSKKCDLAAKLTGLGNVAY 62
 DB 4 HPIIEEVALNTGTEIPYGVKAIPTEAKGSHLIFCHSKKCDLAAKLTGLGNVAY 62

QY 63 RGLDVSIVPTSGDVVW-ATDALMTGFTGDFSDVIDCNTCV 101
 DB 63 RGLDVSIVPTSGDVVW-ATDALMTGFTGDFSDVIDCNTCV 101

Search completed: August 12, 2005, 14:22:47
 Job time : 21.2998 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:03:18 ; Search time 94.231 Seconds
(without alignments)
554.298 Million cell updates/sec

Title: US-09-758-308-2

Perfect score: 536

Sequence: 1 VPHNIEBVALSNTGIEIPFY.....LMTGTFGDFSDVIDCNTCVT 102

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	525.5	98.0	1186	2	Q81755 hepatitis c
2	525.5	98.0	2284	2	Q81817 hepatitis c
3	525.5	98.0	3010	1	P0LG HCVJA
4	525.5	98.0	3010	2	P89966 hepatitis c
5	519.5	96.9	3010	2	Q9J3H5
6	518.5	96.7	3010	2	Q9DTE9
7	517.5	96.5	659	2	Q68K36
8	517.5	96.5	659	2	Q68K57
9	517.5	96.5	3010	2	Q9J3G8
10	517.5	96.5	3011	2	Q68CJ5
11	516.5	96.4	659	2	Q68K50
12	516.5	96.4	3010	2	Q8QRL8
13	516.5	96.4	3015	2	Q9WPH5
14	514.5	96.0	3010	1	P0LG HCVJT
15	514.5	96.0	3010	2	Q9J3H7
16	513.5	95.8	3010	2	Q9DTE7
17	512.5	95.6	3010	2	Q81541
18	512.5	95.6	3010	2	Q9WMX2
19	511.5	95.4	659	2	Q68K44
20	511.5	95.4	659	2	Q68K48
21	511.5	95.4	660	2	Q68K46
22	511.5	95.4	3010	2	Q93016
23	511.5	95.4	3010	2	P90195
24	511.5	95.4	3010	2	Q68285
25	511.5	95.4	3010	2	Q807P3
26	511.5	95.4	3010	2	Q81825
27	511.5	95.4	3010	2	Q9J3H1
28	511.5	95.4	3010	2	Q9J3H2
29	511.5	95.4	3010	2	Q9J3H3
30	511.5	95.4	3010	2	Q9J3H4
31	511.5	95.4	3010	2	Q9Q1Y4

32	511.5	95.4	3010	2	Q9QP06	Q9qp06 hepatitis c
33	511.5	95.4	3011	2	Q9QD8	Q9qtd8 hepatitis c
34	510.5	95.2	593	2	Q9QP07	Q9qp07 hepatitis c
35	510.5	95.2	3010	2	P90194	P90194 hepatitis c
36	510.5	95.2	3010	2	Q9DTE4	Q9dte4 hepatitis c
37	510.5	95.2	3010	2	Q9J3H9	Q9j3h9 hepatitis c
38	510.5	95.2	3010	2	Q9QP61	Q9qp61 hepatitis c
39	510.5	95.2	3013	2	Q9J3H4	Q9j3h4 hepatitis c
40	510.5	95.2	3014	2	Q86614	Q86614 hepatitis c
41	509.5	95.1	661	2	Q68K52	Q68k52 hepatitis c
42	509.5	95.1	3010	2	Q89796	Q89796 hepatitis c
43	509.5	95.1	3010	2	Q9J3F9	Q9j3f9 hepatitis c
44	509.5	95.1	3010	2	Q9J3G4	Q9j3g4 hepatitis c
45	509.5	95.1	3010	2	Q9J3G7	Q9j3g7 hepatitis c

ALIGNMENTS

RESULT 1

Q81755	PRELIMINARY;	PRT; 1186 AA.
ID	Q81755	
AC	Q81755;	
DT	01-NOV-1996 (TREMREL. 01, Created)	
DT	01-NOV-1996 (TREMREL. 01, Last sequence update)	
DT	01-MAR-2004 (TREMREL. 26, Last annotation update)	
DE	Precursor polypeptide (Fragment).	
OS	Hepatitis C virus.	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;	
OC	Hepacivirus.	
ON	NCBI TaxID=111103;	
OX	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=92044440; PubMed=1658196;	
RA	Okamoto H., Kurai K., Iizuka H., Machida A.,	
RA	Miyakawa Y., Mayumi M.;	
RT	"Nucleotide sequences of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions."	
RT	J. Gen. Virol. 72:2697-2704(1991).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=92230232; PubMed=1314459;	
RA	Okamoto H., Kurai K., Okada S., Yamamoto K., Iizuka H., Tanaka T.,	
RA	Fukuda S., Tsuda F., Mishiro S.;	
RT	"Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes."	
RT	Virology 188:331-341(1992).	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=91088550; PubMed=2175903;	
RA	Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,	
RA	Sugimura T., Shimotohno K.;	
RT	"Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis."	
RT	Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).	
RL	[4]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=91140698; PubMed=1847440;	
RA	Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J.,	
RA	Onishi E., Andoh T., Yoshida I., Okayama H.;	
RT	"Structure and organization of the hepatitis C virus genome isolated from human carriers."	
RT	J. Virol. 65:1105-1113(1991).	
RN	[5]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=91172826; PubMed=1848704;	
RA	Choo Q., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,	
RA	Gallegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,	
RA	Bradley D.W., Kuo G., Houghton M.;	
RT	"Genetic organization and diversity of the hepatitis C virus.";	
RT	Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455(1991).	

```

RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=92230206; PubMed=1314449;
RA Chen P., Lin K., Tai K., Liu P., Lin C., Chen D.;
RT "The Taiwanese hepatitis C virus genome: sequence determination and
RL mapping the 5' termini of viral genomic and antigenomic RNA.";
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=93323208; PubMed=8392606;
RA Hijikata M., Mizushima H., Akagi T., Mori S., Kakiuchi N., Kato N.,
RT Tanaka T., Kimura K., Shimotohno K.;
RT "Two distinct proteinase activities required for the processing of a
RL putative structural precursor protein of hepatitis C virus.";
RN [8]
RP SEQUENCE FROM N.A.
RX MEDLINE=94068484; PubMed=7504283;
RA Hijikata M., Mizushima H., Tanji Y., Komada Y., Hirowatari Y.,
RT Akagi T., Kimura K., Shimotohno K.;
RT "Proteolytic processing and membrane association of putative
RL nonstructural proteins of hepatitis C virus.";
RN [9]
RP SEQUENCE FROM N.A.
RX MEDLINE=94333810; PubMed=8056334; DOI=10.1016/0378-1119(94)90008-6;
RA Tanji Y., Hijikata M., Hirowatari Y., Shimotohno K.;
RT "Identification of the domain required for trans-cleavage activity of
RL hepatitis C viral serine proteinase.";
RN [10]
RP SEQUENCE FROM N.A.
RX MEDLINE=95056078; PubMed=7966638;
RA Tanji Y., Hijikata M., Hirowatari Y., Shimotohno K.;
RT "Hepatitis C virus polyprotein processing: kinetics and mutagenic
RL analysis of serine proteinase-dependent cleavage.";
RN [11]
RP SEQUENCE FROM N.A.
RX MEDLINE=95155683; PubMed=7853491;
RA Tanji Y., Hijikata M., Satoh S., Kaneko T., Shimotohno K.;
RT "Hepatitis C virus-encoded nonstructural protein NS4A has versatile
RL functions in viral protein processing.";
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE=951575-1581(1995).
RA J. Virol. 69:1575-1581(1995).
RT EMBL; D16435; BAA03905.1; -.
RL PIR; A61196; A61196.
DR PIR; P00246; P00246.
DR PIR; P00329; P00329.
DR HSP; Q81755; LDXP.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept Ser Cys.
DR InterPro; IPR002518; Pept_U39_HCV NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_FSVir.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Nonstructural protein; Polyprotein; Signal.
FT SIGNAL 4 20

```



```
FT CARBOHYD 2041 2041 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 2077 2077 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 2240 2240 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 2788 2788 N-linked (GlcNAc... ) (Potential).
SQ SEQUENCE 3010 AA; 327017 MW; AA993794F46DB185 CRC64;

Query Match          98.0%; Score 525.5; DB 1; Length 3010;
Best Local Similarity 99.0%; Pred. No. 7.5e-45;
Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 VPHNIEVALSNTGEIPFYKAIPIEAIKGRHLIFCHSKKKCDELAAKLTGLGLNAVA 60
Db 1357 VPHNIEVALSNTGEIPFYKAIPIEAIKGRHLIFCHSKKKCDELAAKLTGLGLNAVA 1416

Qy 61 YYRGLDVSIVPTSGD-VVATDALTMTGTFDSDVIDNCVCT 102
Db 1417 YYRGLDVSIVPTSGD-VVATDALTMTGTFDSDVIDNCVCT 1459

RESULT 4
P89966
ID P89966 PRELIMINARY; PRT; 3010 AA.
AC P89966;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE RNA for polyprotein, complete cds.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP SEQUENCE FROM N.A.
RC STRAIN=type 1b;
RA TANAKA T.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=type 1b;
RA Tanaka T.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D89872; BAA14035.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; Q81755; 1DXP.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0002236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transmembrane; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV NS5b.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
```

```
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral_RGRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327024 MW; E075BD9CFD8D1261 CRC64;

Query Match          98.0%; Score 525.5; DB 2; Length 3010;
Best Local Similarity 99.0%; Pred. No. 7.5e-45;
Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 VPHNIEVALSNTGEIPFYKAIPIEAIKGRHLIFCHSKKKCDELAAKLTGLGLNAVA 60
Db 1357 VPHNIEVALSNTGEIPFYKAIPIEAIKGRHLIFCHSKKKCDELAAKLTGLGLNAVA 1416

Qy 61 YYRGLDVSIVPTSGD-VVATDALTMTGTFDSDVIDNCVCT 102
Db 1417 YYRGLDVSIVPTSGD-VVATDALTMTGTFDSDVIDNCVCT 1459

RESULT 5
Q9J3H5
ID Q9J3H5 PRELIMINARY; PRT; 3010 AA.
AC Q9J3H5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP SEQUENCE FROM N.A.
RC STRAIN=MD17;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF207758; AAF65948.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0254; PQ0254.
DR PIR; PS0329; PS0329.
DR HSP; Q8UY51; 1CW.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0002236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transmembrane; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001545; DEAD/DEAH_N.
```

```

DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002519; HCV core.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Pept Ser Cys.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; Viral RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 326803 MW; 9FEE3D1B93B7AA4B CRC64;

Query Match 96.9%; Score 519.5; DB 2; Length 3010;
Best Local Similarity 97.1%; Pred. No. 3.1e-44;
Matches 100; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 VHPNIEEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKKDELAAKLTGLGNVA 60
Db 1357 VHPNIEEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKKDELAAKLTGLGNVA 1416

QY 61 YRGLDVSVIPSGD-VVATDALTMTGTFGDPDSVIDNCVT 102
Db 1417 YRGLDVSVIPSGD-VVATDALTMTGTFGDPDSVIDNCVT 1459

RESULT 6
Q9DTE9 PRELIMINARY; PRT; 3010 AA.
AC Q9DTE9;
DT 01-WAR-2001 (TrEMBLrel. 16, Created)
DT 01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA Mishihiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
with hepatocellular carcinoma: the 'progression score' revisited.";
RL Hepatol. Res. 20:161-171(2001).
DR EMBL; AB049088; BAB18801.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PQ0329; PQ0329.
DR HSSP; Q8JYS1; 1CWX.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.

```

```

DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR00345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Pept Ser Cys.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; Viral RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 326780 MW; 668CFEAF5FEC3658 CRC64;

Query Match 96.7%; Score 518.5; DB 2; Length 3010;
Best Local Similarity 96.1%; Pred. No. 3.9e-44;
Matches 99; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 1 VHPNIEEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKKDELAAKLTGLGNVA 60
Db 1357 VHPNIEEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKKDELAAKLTGLGNVA 1416

QY 61 YRGLDVSVIPSGD-VVATDALTMTGTFGDPDSVIDNCVT 102
Db 1417 YRGLDVSVIPSGD-VVATDALTMTGTFGDPDSVIDNCVT 1459

RESULT 7
Q68K36 PRELIMINARY; PRT; 659 AA.
AC Q68K36;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,

```



```

RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RL HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685636; AAT94277.1; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70411 MW; 2FCC6D044A67324E1 CRC64;

Query Match 96.5%; Score 517.5; DB 2; Length 659;
Best Local Similarity 97.1%; Pred. No. 1.1e-44;
Matches 100; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 1 VPHNIEEVALSNTEGIEPFYKAIPIEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60
Db 252 VPHNIEEVALSNVEGIEPFYKAIPIESIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 311
Qy 61 YYRGLDVSIVPTSGD-VVVAIDALMTGFTGDFDSVIDNCVT 102
Db 312 YYRGLDVSIVPTSGDVVVVAIDALMTGFTGDFDSVIDNCVT 354

RESULT 8
Q68K57 PRELIMINARY; PRT; 659 AA.
AC Q68K57;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_
RP SEQUENCE FROM N.A.
RA Parker M., Larson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RL HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685592; AAT94256.1; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70411 MW; 2FCC6D044A67324E1 CRC64;

Query Match 96.5%; Score 517.5; DB 2; Length 659;
Best Local Similarity 97.1%; Pred. No. 1.1e-44;
Matches 100; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

```

```

Qy 1 VPHNIEEVALSNTEGIEPFYKAIPIEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60
Db 252 VPHNIEEVALSNVEGIEPFYKAIPIESIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 311
Qy 61 YYRGLDVSIVPTSGD-VVVAIDALMTGFTGDFDSVIDNCVT 102
Db 312 YYRGLDVSIVPTSGDVVVVAIDALMTGFTGDFDSVIDNCVT 354

RESULT 9
Q9J3G8 PRELIMINARY; PRT; 3010 AA.
AC Q9J3G8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=MD24;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF207765; AAF65955.1; -.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSSP; O8JYS1; 1CWX.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_neme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR003519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral_RdRP; 1.
DR SMART; SM00487; DEXDC; 1.

```

DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 326863 MW; E0BEB3528215457C CRC64;
Query Match 96.5%; Score 517.5; DB 2; Length 3010;
Best Local Similarity 97.1%; Pred. No. 5e-44;
Matches 100; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
QY 1 VPHNIEEVALSNTGEIPYKAIPIEAIKGRHLIFCHSKKCDLAALKLGLGNVA 60
|||
Db 1357 VPHNIEEVALSNTGEIPYKAIPIEAIKGRHLIFCHSKKCDLAALKLGLGNVA 1416
|||
QY 61 YRGLDVSVIPSGD-VVATDALTMTGTFGDFSDSDVDCNTCVT 102
|||
Db 1417 YRGLDVSVIPSGD-VVATDALTMTGTFGDFSDSDVDCNTCVT 1459
|||
RESULT 10
Q6SCJ5 PRELIMINARY; PRT; 3011 AA.
AC Q6SCJ5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HCV-S;
RA Yuan Z., Mao H., Hu Y., Lan S., Wang H.;
RT "Construction of full-length complementary DNA of Hepatitis C Virus genome.";
RL Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi 0:0-0(2004).
DR EMBL; AY460204; AR22408.1; -
DR HSSP; P26633; IJXP.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NSI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39 HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NSI; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF01006; HCV_NS4a; 1.

DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; Viral_RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3011 AA; 326743 MW; 9216EF485C7E9036 CRC64;
Query Match 96.5%; Score 517.5; DB 2; Length 3011;
Best Local Similarity 97.1%; Pred. No. 5e-44;
Matches 100; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
QY 1 VPHNIEEVALSNTGEIPYKAIPIEAIKGRHLIFCHSKKCDLAALKLGLGNVA 60
|||
Db 1357 VPHNIEEVALSNTGEIPYKAIPIEAIKGRHLIFCHSKKCDLAALKLGLGNVA 1416
|||
QY 61 YRGLDVSVIPSGD-VVATDALTMTGTFGDFSDSDVDCNTCVT 102
|||
Db 1417 YRGLDVSVIPSGD-VVATDALTMTGTFGDFSDSDVDCNTCVT 1459
|||
RESULT 11
Q68K50 PRELIMINARY; PRT; 659 AA.
AC Q68K50;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in HCV/HIV Coinfected Subjects."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685606; AAT94263.1; -
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW Polyprotein.
KW NON_TER 1
FT NON_TER 659
FT SEQUENCE 659 AA; 70519 MW; B48CCE1A0C6F9EDD CRC64;
Query Match 96.4%; Score 516.5; DB 2; Length 659;
Best Local Similarity 96.1%; Pred. No. 1.4e-44;
Matches 99; Conservative 3; Mismatches 0; Indels 1; Gaps 1;
QY 1 VPHNIEEVALSNTGEIPYKAIPIEAIKGRHLIFCHSKKCDLAALKLGLGNVA 60
|||
Db 252 VPHNIEEVALSNTGEIPYKAIPIEAIKGRHLIFCHSKKCDLAALKLGLGNVA 311
|||
QY 61 YRGLDVSVIPSGD-VVATDALTMTGTFGDFSDSDVDCNTCVT 102
|||
Db 312 YRGLDVSVIPSGD-VVATDALTMTGTFGDFSDSDVDCNTCVT 354
|||
RESULT 12
Q8QRL8

```

ID Q8QRL8 PRELIMINARY; PRT; 3010 AA.
AC Q8QRL8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus; Hepatitis C virus type 1.
OX NCBI_TaxID=31647;
RN [1]
RP SEQUENCE FROM N.A.
RA Yildiz E., Oztan A., Akkiz H., Ozturk M.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF483269; AAL91977.1; -.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CWK.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0006508; F: structural molecule activity; IEA.
DR GO; GO:0006350; F: proteolysis and peptidolysis; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA pol_DS_PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 326932 MW; 3BE8CA9C861814CB CRC64;

Query Match 96.4%; Score 516.5; DB 2; Length 3010;
Best Local Similarity 96.1%; Pred. No. 6.3e-44;
Matches 99; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEEVALNTGTEIPYKAIPTEAIKGGRHILFCHSKKKDELAAKLTGLGLNAVA 60
Db 1357 VPHNIEEVALNTGTEIPYKAIPVEVIKGGRHILFCHSKKKDELAAKLTGLGLNAVA 1416
```

```

QY 61 YYRGLDVSVIPTSGD-VVVAATDALMTGFTGDFDVIDCNCVT 102
Db 1417 YYRGLDVSVIPTSGDVVVVAATDALMTGFTGDFDVIDCNCVT 1459

RESULT 13
Q9WPH5 PRELIMINARY; PRT; 3015 AA.
AC Q9WPH5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=HCV-N;
RC MEDLINE=99315771; PubMed=10385673; DOI=10.1002/hep.510300137;
RA Beard M.R., Abell G., Honda M., Carroll A., Gartland M., Clarke B.,
RA Suzuki K., Lanford R., Sangar D.V., Lemon S.M.;
RA Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF139594; AAD44718.2; -.
DR HSSP; Q8JYS1; 1CWK.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0006508; F: structural molecule activity; IEA.
DR GO; GO:0006350; F: proteolysis and peptidolysis; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA pol_DS_PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
```

DR Pfam: PF00271; Helicase C; 1.
 DR Pfam: PF00998; Viral RdRp; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR PROSITE: PS00190; CYTOCHROME C; UNKNOWN 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 SQ SEQUENCE 3015 AA; A86AE71196578BE3 CRC64;

Query Match 96.4%; Score 516.5; DB 2; Length 3015;
 Best Local Similarity 96.1%; Pred. No. 6.3e-44;
 Matches 99; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGEIPYKAIPIEALKGRHLIFCHSKKCDLAAKLTGLGINAVA 60
 DB 1358 VPHNIEEVALSNTGEIPYKAIPIEALKGRHLIFCHSKKCDLAAKLTGLGINAVA 1417
 QY 61 YYRGLDVSVIPSGD-VVWATDALMTGFTGDPDSVIDNCVT 102
 DB 1418 YYRGLDVSVIPSGDVVVWATDALMTGVTGDPDSVIDNCVT 1460

RESULT 14
 POLG_HCVJT
 ID POLG_HCVJT STANDARD; PRT; 3010 AA.
 AC Q00269;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 45, Last annotation update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
 DE Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
 DE (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
 DE (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
 DE (EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
 DE NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
 DE NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
 OS Hepatitis C virus (isolate HC-JT) (HCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=31642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92295714; PubMed=1318627; DOI=10.1016/0168-1702(92)90066-1;
 RA Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J., Nakazawa T.,
 RA Hijikata M., Ishimura Y., Shimotohno K.;
 RT "Molecular cloning of hepatitis C virus genome from a single Japanese
 RT carrier: sequence variation within the same individual and among
 RT infected individuals."
 RL Virus Res. 23:39-53(1992).
 CC -1- FUNCTION: The small proteins NS2A, NS2B, NS4A and NS4B are
 CC hydrophobic, suggesting a possible membrane-related function. NS3
 CC and NS5 may play a role in the viral RNA replication.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
 CC precursor polyprotein, commonly with Asp or Glu in the P6
 CC position, Cys or Thr in P1 and Ser or Ala in P1'.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC [RNA] (N).
 CC -1- SUBUNIT: The virion of this virus is a nucleocapsid covered by a
 CC lipoprotein envelope. The envelope consists of two proteins:
 CC protein M and glycoprotein E. The nucleocapsid is a complex of
 CC protein C and mRNA.
 CC -1- SIMILARITY: Contains 1 peptidase S29 domain.
 CC -1- SIMILARITY: Contains 1 peptidase U39 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D11168; BAA01943.1; --
 CC PIR: A45573; A45573.

DR HSP; P26663; IJXP.
 DR MEROPS: S29.001; --
 DR MEROPS: U39.001; --
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR002522; HCV capsid.
 DR InterPro: IPR002521; HCV_core.
 DR InterPro: IPR002519; HCV env.
 DR InterPro: IPR002531; HCV NS1.
 DR InterPro: IPR000745; HCV NS4a.
 DR InterPro: IPR001490; HCV NS4b.
 DR InterPro: IPR002868; HCV NS5a.
 DR InterPro: IPR002166; HCV RdRp.
 DR InterPro: IPR001650; Helicase C.
 DR InterPro: IPR009003; Pept Ser Cys.
 DR InterPro: IPR002518; Pept U39 HCV NS2.
 DR InterPro: IPR004109; Peptidase S29.
 DR InterPro: IPR007095; RNA pol_D5_P5.
 DR InterPro: IPR007094; RNA pol_P5vir.
 DR Pfam: PF01543; HCV capsid; 1.
 DR Pfam: PF01542; HCV core; 1.
 DR Pfam: PF01539; HCV env; 1.
 DR Pfam: PF01560; HCV NS1; 1.
 DR Pfam: PF01538; HCV NS2; 1.
 DR Pfam: PF02907; HCV NS3; 1.
 DR Pfam: PF01006; HCV NS4a; 1.
 DR Pfam: PF01001; HCV NS4b; 1.
 DR Pfam: PF01506; HCV NS5a; 1.
 DR Pfam: PF00271; Helicase C; 1.
 DR Pfam: PF00998; Viral RdRp; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR ATP-binding; Coat protein; Core protein; Envelope protein;
 KW Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
 KW RNA-directed RNA polymerase; Serine protease; Transferase;
 KW Transmembrane.
 FT INIT_MET 1
 FT CHAIN 1 115
 FT CHAIN 116 191
 FT CHAIN 192 383
 FT CHAIN 384 729
 FT CHAIN 730 1006
 FT CHAIN 1007 1615
 FT CHAIN 1616 1862
 FT CHAIN 1863 2013
 FT CHAIN 2014 3010
 FT TRANSMEM 347 389
 FT ACT SITE 1083 1083
 FT ACT SITE 1107 1107
 FT ACT SITE 1165 1165
 FT NP_BIND 1230 1237
 FT SITE 1316 1319
 FT CARBOHYD 136 136
 FT CARBOHYD 209 209
 FT CARBOHYD 234 234
 FT CARBOHYD 250 250
 FT CARBOHYD 305 305
 FT CARBOHYD 417 417
 FT CARBOHYD 423 423
 FT CARBOHYD 430 430
 FT CARBOHYD 448 448
 FT CARBOHYD 532 532
 FT CARBOHYD 540 540
 FT CARBOHYD 556 556
 FT CARBOHYD 576 576
 FT CARBOHYD 623 623
 FT CARBOHYD 645 645
 FT CARBOHYD 2041 2041
 FT CARBOHYD 2077 2077
 FT CARBOHYD 2240 2240
 FT CARBOHYD 2529 2529
 FT CARBOHYD 2788 2788
 FT SEQUENCE 3010 AA; 326573 MW; 94A1C77435D642BB CRC64;

```
Query Match          96.0%; Score 514.5; DB 1; Length 3010;
Best Local Similarity 95.1%; Pred. No. 1e-43;
Matches 98; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGEIPFYGKAIPLEAIGRHLIFCHSKKCKDELAALKLGLGLNAVA 60
    |||||
DB 1357 VPHNIEEVALSNTGEIPFYGKAIPLEAIGRHLIFCHSKKCKDELAALKLGLGLNAVA 1416

QY 61 YYRGLDVSIVPTSGDVV-VATDALMTGTGDFSDVIDCNCVT 102
    |||||
DB 1417 YYRGLDVSIVPTSGDVV-VATDALMTGTGDFSDVIDCNCVT 1459

RESULT 15
ID Q9J3H7 PRELIMINARY; PRT; 3010 AA.
AC Q9J3H7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MD15;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF207756; AAF65946.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; O8JYS1; 1CWX.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NSI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NSI; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
```

```
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; Viral_RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327366 MW; D8653F7317FFA106 CRC64;

Query Match          96.0%; Score 514.5; DB 2; Length 3010;
Best Local Similarity 96.1%; Pred. No. 1e-43;
Matches 99; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGEIPFYGKAIPLEAIGRHLIFCHSKKCKDELAALKLGLGLNAVA 60
    |||||
DB 1357 VPHNIEEVALSNTGEIPFYGKAIPLEAIGRHLIFCHSKKCKDELAALKLGLGLNAVA 1416

QY 61 YYRGLDVSIVPTSGD-VVATDALMTGTGDFSDVIDCNCVT 102
    |||||
DB 1417 YYRGLDVSIVPTSGD-VVATDALMTGTGDFSDVIDCNCVT 1459
```

Search completed: August 12, 2005, 14:21:19
Job time : 96.231 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:00:18 ; Search time 107.012 Seconds
(without alignments)
368.645 Million cell updates/sec

Title: US-09-758-308-2

Perfect score: 536

Sequence: 1 VPHNIEEVALSNTGPIPFY.....LMTGFTGDFSDVIDNCVT 102

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 16Dec04:*

- 1: geneseqp1980as:*
- 2: geneseqp1990as:*
- 3: geneseqp2000as:*
- 4: geneseqp2001as:*
- 5: geneseqp2002as:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	525.5	98.0	211	2	AAR52736 HCV anti
2	525.5	98.0	226	2	AAR25860 HCV poly
3	525.5	98.0	226	2	AAW41738 Hepatitis
4	525.5	98.0	252	2	AAR25882 HK7. 9/20
5	525.5	98.0	631	2	AAW26160 Serine pr
6	525.5	98.0	631	2	AAW14354 Hepatitis
7	525.5	98.0	631	2	AAV15806 HCV strai
8	525.5	98.0	631	7	ADL18158 Hepatitis
9	525.5	98.0	916	2	AAR82693 HCV parti
10	525.5	98.0	923	2	AAR82696 HCV parti
11	525.5	98.0	3010	2	AAR68622 HCV prote
12	525.5	98.0	3010	2	AAR68864 Hepatitis
13	525.5	98.0	3010	2	AAR82694 Partial H
14	525.5	98.0	3010	7	ADF88597 Hepatitis
15	521.5	97.3	101	4	AAB31696 Antigenic
16	517.5	96.5	1736	4	AAB36932 Hepatitis
17	516.5	96.4	632	8	ADO36213 Hepatitis
18	516.5	96.4	632	8	ADO79387 Hepatitis
19	514.5	96.0	2280	8	ADI95303 OSPF-rela
20	513.5	95.8	667	2	AAV17893 HCV NS4a
21	512.5	95.6	1985	5	AAOI18001 Hepatitis
22	512.5	95.6	1985	5	AAEI5729 Hepatitis
23	512.5	95.6	1985	5	AAEI5731 Hepatitis
24	512.5	95.6	1985	5	AAEI5720 Hepatitis
25	512.5	95.6	1985	5	AAEI5717 Hepatitis

ALIGNMENTS

RESULT 1

AAR52736
ID AAR52736 standard; protein; 211 AA.

AC AAR52736;

DT 31-JAN-1995 (first entry)

DE HCV antigen.

KW Hepatitis C virus; HCV; antigen; diagnosis; reagent; agglutination.

OS Synthetic.

PN JP06102273-A.

PD 15-APR-1994.

PF 18-SEP-1992; 92JP-00250027.

PR 18-SEP-1992; 92JP-00250027.

PA (TOKU) TOKUYAMA SODA KK.

WPI: 1994-161280/20.

N-PSDB; AAQ62689.

Immunological agglutination reagent for the diagnosis of hepatitis C - comprising hepatitis C virus antigen polypeptide.

Claim 2-3; Page 16-17; 18pp; Japanese.

An new immunological agglutination reaction reagent for the diagnosis of hepatitis C uses a HCV antigen polypeptide subjected to heat-treatment. The HCV antigen active polypeptide contains one of the polypeptides given in AAR52735-38

SQ Sequence 211 AA;

Query Match 98.0%; Score 525.5; DB 2; Length 211;

Best Local Similarity 99.0%; Pred. No. 1.1e-57;

Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGPIPFYKAIPIEATKGGHILFCHSKKCDLAAKLTGLGNVA 60

DB 35 VPHNIEEVALSNTGPIPFYKAIPIEATKGGHILFCHSKKCDLAAKLTGLGNVA 94

QY 61 YYRGLDVSIVTSGD-VVATDALTMTGFTGDFSDVIDNCVT 102

26 512.5 95.6 1985 5 AAE15727 Hepatitis
27 512.5 95.6 1985 5 AAE15728 Hepatitis
28 512.5 95.6 1985 5 AAE15722 Hepatitis
29 512.5 95.6 1985 5 AAE15730 Hepatitis
30 512.5 95.6 1985 8 ADJ57846 HCV repli
31 512.5 95.6 2063 7 ADG67963 Hepatitis
32 512.5 95.6 2201 5 ABG30601 Hepatitis
33 512.5 95.6 2201 5 ABG30584 Hepatitis
34 512.5 95.6 2201 5 ABG30591 Hepatitis
35 512.5 95.6 2201 5 ABG30600 Hepatitis
36 512.5 95.6 2201 5 ABG30581 Hepatitis
37 512.5 95.6 2201 5 ABG30586 Hepatitis
38 512.5 95.6 2201 5 ABG30593 Hepatitis
39 512.5 95.6 2201 5 ABG30582 Hepatitis
40 512.5 95.6 2201 5 ABG30580 Hepatitis
41 512.5 95.6 2201 5 ABG30602 Hepatitis
42 512.5 95.6 2201 5 ABG30587 Hepatitis
43 512.5 95.6 2201 5 ABG30589 Hepatitis
44 512.5 95.6 2201 5 ABG30599 Hepatitis
45 512.5 95.6 2201 5 ABG30585 Hepatitis

```
Db      95 YRGLDVSVIPTSGD VVVVATDALMTGFTGDFDSDVIDNCVT 137
|||||
RESULT 2
ID AAR25860 standard; protein; 226 AA.
XX
AC AAR25860;
XX
DT 21-JAN-1993 (first entry)
XX
DE HCV polypeptide 7.
XX
KW Recombinant vector; E. coli; diagnostic; reagent; type C hepatitis.
XX
OS Hepatitis C virus.
XX
PN JP04179482-A.
XX
PD 26-JUN-1992.
XX
PF 11-NOV-1990; 90JP-00304417.
XX
PR 11-NOV-1990; 90JP-00304417.
XX
PA (TOKU ) TOKUYAMA SODA KK.
XX
DR WPI; 1992-263663/32.
XX
DR N-PSDB; AAQ26987.
XX
Hepatitis C virus antigen expressed as recombinant in E.coli - useful for
PT diagnosis of hepatitis C virus infection.
XX
PS Claim 1; Page 3; 66pp; Japanese.
XX
CC The sequences given in AAR25854-74 are hepatitis C virus proteins. The
CC genes encoding these proteins can each be used to prepare recombinant
CC vectors by ligating the gene of interest in to a vector to be expressed
CC in E. coli. These polypeptides are useful as diagnostic reagents for type
CC C hepatitis and they may be produced efficiently by recombinant methods
XX
SQ Sequence 226 AA;
Query Match 98.0%; Score 525.5; DB 2; Length 226;
Best Local Similarity 99.0%; Pred. No. 1.2e-57;
Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 VPHPNIEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKKCDELAALKLTGLGLNAVA 60
DB 50 VPHPNIEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKKCDELAALKLTGLGLNAVA 109
QY 61 YRGLDVSVIPTSGD-VVVATDALMTGFTGDFDSDVIDNCVT 102
DB 110 YRGLDVSVIPTSGD VVVVATDALMTGFTGDFDSDVIDNCVT 152
|||||
RESULT 4
AAR25882
ID AAR25882 standard; protein; 252 AA.
XX
AC AAR25882;
XX
DT 09-SEP-2004 (revised)
DT 21-JAN-1993 (first entry)
XX
DE HK7.
XX
KW Recombinant vector; E. coli; diagnostic; reagent; type C hepatitis.
XX
OS Hepatitis C virus.
OS Unidentified.
XX
Key Location/Qualifiers
FH Protein 13..238
FT /note= "Sequence AAR25860"
XX
PN JP04179482-A.
XX
PD 26-JUN-1992.
XX
PF 11-NOV-1990; 90JP-00304417.
XX
PR 11-NOV-1990; 90JP-00304417.
XX
PA (TOKU ) TOKUYAMA SODA KK.
XX
DR WPI; 1992-263663/32.
XX
DR N-PSDB; AAQ27009.
XX
Hepatitis C virus antigen expressed as recombinant in E.coli - useful for
PT diagnosis of hepatitis C virus infection.
XX
PS Disclosure; Fig 8; 66pp; Japanese.
```


XX The sequences given in AAR25876-95 are encoded by the claimed hepatitis C
 CC virus genes of the invention which have been inserted into an E. coli
 CC vector. These polypeptides are useful as diagnostic reagents for type C
 CC hepatitis and they may be produced efficiently by recombinant DNA
 CC techniques

CC Revised record issued on 09-SEP-2004 : Correction to feature table key

XX Sequence 252 AA;

Query Match 98.0%; Score 525.5; DB 2; Length 252;
 Best Local Similarity 99.0%; Pred. No. 1.4e-57;
 Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHPNIEEVALSNTGEIPFYGKAIPTEIAIKGRHLIFCHSKKKCDELAAKLTGLGLNAVA 60
 |||||
 DB 62 VPHPNIEEVALSNTGEIPFYGKAIPTEIAIKGRHLIFCHSKKKCDELAAKLTGLGLNAVA 121
 |||||
 QY 61 YRGLDVSIVPTSGD-VVATDALTMTGFTGDFSDVIDNCVT 102
 |||||
 DB 122 YRGLDVSIVPTSGD-VVATDALTMTGFTGDFSDVIDNCVT 164
 |||||

RESULT 5

AAW26160
 ID AAW26160 standard; peptide; 631 AA.

XX AAW26160;

XX 24-NOV-1997 (first entry)

DE Serine protease NS3 region.

XX Serine protease; NS3 region; HCV; hepatitis C virus; monoclonal antibody;
 KW IgG; immunocyte; myeloma cell; inhibitor; neutralising agent.

XX Hepatitis C virus.

XX JP09206076-A.

XX 12-AUG-1997.

XX 06-FEB-1996; 96JP-00020321.

XX 06-FEB-1996; 96JP-00020321.

XX (NIHA) JAPAN ENERGY CORP.

XX WPI; 1997-451976/42.

XX N-PSDB; AAT80095.

XX Monoclonal antibody against hepatitis C virus serine protease - useful
 PT for diagnosis of HCV infection.

XX Claim 2; Page 14-17; 22pp; Japanese.

XX This sequence represents the NS3 region of the hepatitis C virus (HCV)
 CC serine protease. The fragment of this sequence represented in AAW26159 is
 CC recognised by the monoclonal antibody of the invention. The monoclonal
 CC antibody belongs to the IgG class of antibodies. The antibody is produced
 CC by fusion cells formed between immunocytes from a HCV infected mammal,
 CC and myeloma cells from the same mammal. The antibody is used for
 CC diagnosis of HCV infection. As this monoclonal antibody can bind to a
 CC certain chain in serine protease and inhibits the enzyme specifically, it
 CC can also be used as an agent for neutralising the activity of serine
 CC protease

XX Sequence 631 AA;

Query Match 98.0%; Score 525.5; DB 2; Length 631;
 Best Local Similarity 99.0%; Pred. No. 4.8e-57;
 Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHPNIEEVALSNTGEIPFYGKAIPTEIAIKGRHLIFCHSKKKCDELAAKLTGLGLNAVA 60
 |||||
 DB 331 VPHPNIEEVALSNTGEIPFYGKAIPTEIAIKGRHLIFCHSKKKCDELAAKLTGLGLNAVA 390
 |||||
 QY 61 YRGLDVSIVPTSGD-VVATDALTMTGFTGDFSDVIDNCVT 102
 |||||
 DB 391 YRGLDVSIVPTSGD-VVATDALTMTGFTGDFSDVIDNCVT 433
 |||||

RESULT 6

AAI14354
 ID AAI14354 standard; protein; 631 AA.

XX AAI14354;

XX 17-AUG-1999 (first entry)

XX Hepatitis C virus Ser/Thr protease amino acid sequence.

XX Complementarity determining region; CDR; monoclonal antibody; MAb;
 KW hepatitis C virus; HCV; protease.

XX Hepatitis C virus.

XX JP11127861-A.

XX 18-MAY-1999.

XX 29-OCT-1997; 97JP-00297451.

XX 29-OCT-1997; 97JP-00297451.

XX (NIHA) JAPAN ENERGY CORP.

XX WPI; 1999-350322/30.

XX N-PSDB; AAX57785.

XX Neutralized antibody partial peptide derived from hepatitis C virus -
 PT useful for inhibiting Hepatitis C Virus (HCV) serine protease activity.

XX Disclosure; Page 14-16; 32pp; Japanese.

XX This sequence represents the amino acid sequence of a Ser/Thr protease
 CC from hepatitis C virus. The invention relates to the use of partial
 CC peptides (AAI14348-YI4353) from the anti-HCV neutralising MAb 8D4 for
 CC inhibiting HCV serine protease activity

XX Sequence 631 AA;

Query Match 98.0%; Score 525.5; DB 2; Length 631;
 Best Local Similarity 99.0%; Pred. No. 4.8e-57;
 Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHPNIEEVALSNTGEIPFYGKAIPTEIAIKGRHLIFCHSKKKCDELAAKLTGLGLNAVA 60
 |||||
 DB 331 VPHPNIEEVALSNTGEIPFYGKAIPTEIAIKGRHLIFCHSKKKCDELAAKLTGLGLNAVA 390
 |||||

QY 61 YRGLDVSIVPTSGD-VVATDALTMTGFTGDFSDVIDNCVT 102
 |||||
 DB 391 YRGLDVSIVPTSGD-VVATDALTMTGFTGDFSDVIDNCVT 433
 |||||

RESULT 7

AAI15806

ID AAI15806 standard; protein; 631 AA.

XX AAI15806;

XX 27-JUL-1999 (first entry)

XX HCV strain J antigen sequence.

KW Antigen peptide; HCV J strain; chain peptide; detection; antibody.
 XX Hepatitis C virus.
 XX JP11124398-A.
 PD 11-MAY-1999.
 XX 22-OCT-1997; 97JP-00290165.
 XX 22-OCT-1997; 97JP-00290165.
 XX (NIHA) JAPAN ENERGY CORP.
 XX WPI; 1999-341639/29.
 DR N-PSDB; AAX59785.
 XX New antigen peptide from hepatitis C virus - useful in examination agent
 PT for antibody.
 XX Disclosure; Page 10-13; 25pp; Japanese.
 CC The specification describes an antigen peptide derived from hepatitis C
 CC virus (HCV), which comprises a chain peptide of at least 12 amino acid
 CC residues. The peptide chain comprises at least one amino acid sequence
 CC selected from Gly Trp Pro or AAY15764-66. The peptide is useful in
 CC methods to detect antibodies directed against HCV. The present sequence
 CC represents a HCV strain J antigen
 XX Sequence 631 AA;
 SQ Query Match 98.0%; Score 525.5; DB 2; Length 631;
 Best Local Similarity 99.0%; Pred. No. 4.8e-57;
 Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 VPHPNIEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60
 DB 331 VPHPNIEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 390
 QY 61 YYRGLDVSVIPTSGD-VVATDALTMTGTFGDFSDVIDNCNTCVT 102
 DB 391 YYRGLDVSVIPTSGD-VVATDALTMTGTFGDFSDVIDNCNTCVT 433
 RESULT 8
 ADL18158
 ID ADL18158 standard; protein; 631 AA.
 XX ADL18158;
 XX 06-MAY-2004 (first entry)
 XX Hepatitis C virus NS3 protease protein SEQ ID NO:78.
 XX chimeric protein; signal protein; trafficking signal targeting;
 KW proteolytic cleavage site; protease; protease inhibitor; enzyme.
 XX Hepatitis C virus.
 OS WO2003014381-A1.
 XX 20-FEB-2003.
 XX 08-AUG-2002; 2002WO-KR001515.
 XX 10-AUG-2001; 2001KR-00048123.
 XX (AHRA-) AHRAM BIOSYSTEMS INC.
 XX Hwang I, Kim DH, Lee YJ;
 XX WPI; 2003-256596/25.
 DR N-PSDB; ADL18157.
 DR

XX New chimeric protein, useful for detecting protease inhibitors inside the
 PT cell or tissue.
 XX Disclosure; SEQ ID NO 78; 214pp; English.
 XX The present invention describes a chimeric protein comprising at least
 CC one signal protein that has a trafficking signal targeting to a
 CC subcellular organelle and at least one proteolytic cleavage site for a
 CC protease. The chimeric protein is constructed, so that: (a) the
 CC trafficking signals of all the signal proteins are inactivated by linking
 CC the proteolytic site or a signal masking protein through the proteolytic
 CC site to the N- or C- terminus of the signal proteins, and so the chimeric
 CC protein is present in cytosol; (b) the trafficking signal of at least one
 CC signal protein is activated when the proteolytic cleavage site is cleaved
 CC by the protease, and as a result at least one fragment protein that
 CC includes the activated chimeric protein is transported to a subcellular
 CC organelle; and (c) the chimeric protein is labelled with at least one
 CC fluorescent protein and the position and intensity distribution of the
 CC fluorescent label signal in the cell is altered depending on the cleavage
 CC by the protease. Also described: (1) a recombinant gene comprising a
 CC nucleic acid sequence encoding the chimeric protein which is constructed
 CC to express the chimeric protein in a cell; (2) a cell transformed with
 CC the recombinant gene or vector; (3) analysing the activity of a protease
 CC in vivo; (4) screening protease inhibitors in vivo; (5) a system for
 CC detecting a protease inside a cell; (6) a nucleic acid comprising the
 CC sequence encoding the chimeric protein for detecting protease activity in
 CC a cell; (7) a vector comprising the nucleic acid; (8) a kit for detecting
 CC a protease inside a cell comprising the chimeric protein or the vector;
 CC (9) detecting a protease inside a cell or tissue; and (10) detecting a
 CC protease inhibitor in vivo. The chimeric protein is useful for detecting
 CC protease inhibitors inside the cell or tissue. The present sequence
 CC represents HCV NS3 protease, which is used in the exemplification of the
 CC present invention.
 XX Sequence 631 AA;
 SQ Query Match 98.0%; Score 525.5; DB 7; Length 631;
 Best Local Similarity 99.0%; Pred. No. 4.8e-57;
 Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 VPHPNIEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60
 DB 331 VPHPNIEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 390
 QY 61 YYRGLDVSVIPTSGD-VVATDALTMTGTFGDFSDVIDNCNTCVT 102
 DB 391 YYRGLDVSVIPTSGD-VVATDALTMTGTFGDFSDVIDNCNTCVT 433
 RESULT 9
 AAR82693
 ID AAR82693 standard; protein; 916 AA.
 XX AAR82693;
 XX 16-OCT-2003 (revised)
 DT 11-NOV-1996 (first entry)
 XX HCV partial proteinase.
 XX proteinase; hepatitis C virus; screening; inhibitor; proteolytic;
 KW identification; cleavage.
 XX Hepatitis C virus; Virus.
 XX JP07184648-A.
 XX 25-JUL-1995.
 XX 05-FEB-1993; 93JP-00018854.
 XX 07-FEB-1992; 92JP-00022657.
 PR

PR 18-SEP-1992; 92JP-00249240.
 PR 04-DEC-1992; 92JP-00325303.
 XX (KAEN/) KAENNO K.
 PA (SUMQ) SUMITOMO METAL IND LTD.
 PA (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.
 XX WPI; 1995-287962/38.
 DR N-PSDB; AAT03959.
 XX An HCV proteinase active substance - which has activity as an anti-HCV
 PT agent and can be used to screen for proteinase inhibitors.
 PT
 XX Claim 3; Page 24-27; 52pp; Japanese.
 PS
 XX The present sequence is that of a partial proteinase isolated from
 CC Hepatitis C virus (HCV). The proteinase can be used as an anti-HCV agent.
 CC It can also be used to screen cpds. for their ability to inhibit its
 CC proteolytic activity. In this way proteinase inhibitors can be
 CC identified. (Updated on 16-OCT-2003 to standardise OS field)
 XX
 XX Sequence 916 AA;
 SQ
 Query Match 98.0%; Score 525.5; DB 2; Length 916;
 Best Local Similarity 99.0%; Pred. No. 7.8e-57;
 Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 AC AAR82696;
 XX
 QY 1 VPHNIEVALSNTGEIPFYGKAIPTEIAIKGRHLIFCHSKKCDLAALKLTGLGNAVA 60
 DB 366 VPHNIEVALSNTGEIPFYGKAIPTEIAIKGRHLIFCHSKKCDLAALKLTGLGNAVA 425
 QY 61 YRGLDVSIVPTSGD-VVATDALTMTGTFDSDVIDCNCVT 102
 DB 426 YRGLDVSIVPTSGD-VVATDALTMTGTFDSDVIDCNCVT 468
 RESULT 10
 AAR82696
 ID AAR82696 standard; protein; 923 AA.
 XX
 AC AAR82696;
 XX
 DT 16-OCT-2003 (revised)
 DT 14-NOV-1996 (first entry)
 XX
 XX HCV partial proteinase.
 DE
 XX proteinase; hepatitis C virus; screening; inhibitor; proteolytic;
 KW identification; cleavage.
 KW
 XX Hepatitis C virus; Virus.
 OS
 XX JP07184648-A.
 PN
 XX 25-JUL-1995.
 PD
 XX
 PF 05-FEB-1993; 93JP-00018854.
 XX
 PR 07-FEB-1992; 92JP-00022657.
 PR 18-SEP-1992; 92JP-00249240.
 PR 04-DEC-1992; 92JP-00325303.
 XX
 XX (KAEN/) KAENNO K.
 PA (SUMQ) SUMITOMO METAL IND LTD.
 PA (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.
 XX WPI; 1995-287962/38.
 DR N-PSDB; AAT03985.
 XX
 XX An HCV proteinase active substance - which has activity as an anti-HCV
 PT agent and can be used to screen for proteinase inhibitors.
 PT
 XX Example 2; Page 33-39; 52pp; Japanese.
 PS

XX The present sequence (contg. AAR82693) is that of a partial proteinase
 CC isolated from Hepatitis C virus (HCV). The proteinase can be used as an
 CC anti-HCV agent. It can also be used to screen cpds. for their ability to
 CC inhibit its proteolytic activity. In this way proteinase inhibitors can
 CC be identified. (Updated on 16-OCT-2003 to standardise OS field)
 XX
 XX Sequence 923 AA;
 SQ
 Query Match 98.0%; Score 525.5; DB 2; Length 923;
 Best Local Similarity 99.0%; Pred. No. 7.9e-57;
 Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 VPHNIEVALSNTGEIPFYGKAIPTEIAIKGRHLIFCHSKKCDLAALKLTGLGNAVA 60
 DB 368 VPHNIEVALSNTGEIPFYGKAIPTEIAIKGRHLIFCHSKKCDLAALKLTGLGNAVA 427
 QY 61 YRGLDVSIVPTSGD-VVATDALTMTGTFDSDVIDCNCVT 102
 DB 428 YRGLDVSIVPTSGD-VVATDALTMTGTFDSDVIDCNCVT 470
 RESULT 11
 AAR86822
 ID AAR86822 standard; protein; 3010 AA.
 XX
 AC AAR86822;
 XX
 DT 16-OCT-2003 (revised)
 DT 16-OCT-1995 (first entry)
 XX
 XX HCV protein cleavable with new serine proteinase.
 DE
 XX proteinase; serine; cleavage; hepatitis C virus; HCV.
 KW
 XX Hepatitis C virus; Virus.
 OS
 XX
 FH Key Location/Qualifiers
 FT Cleavage-site 2419..2420
 FT /note= "Serine protease cleavage site"
 XX
 XX JP06315377-A.
 PN
 XX 15-NOV-1994.
 PD
 XX 06-MAY-1993; 93JP-00105666.
 PF
 XX 06-MAY-1993; 93JP-00105666.
 PR
 XX (KAEN/) KAENNO K.
 PA (SUMQ) SUMITOMO METAL IND LTD.
 PA (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.
 XX WPI; 1995-032330/05.
 DR N-PSDB; AAR80498.
 XX
 XX New HCV-originated proteinase active substance - used for site-specific
 PT cleavage by an intermolecular reaction and the purification thereof.
 PT
 XX Disclosure; Page 10-19; 23pp; Japanese.
 PS
 XX This protein from HCV (hepatitis C virus) (encoded by AAR80498) is
 CC cleaved between amino acids 2419 and 2420, by a new serine protease,
 CC contg. the sequence of AAR86821. The proteinase is purified as a fused
 CC product with the dihydrofolate reductase protein by using a methotrexate
 CC column. It can be used for the development of an inhibitor for HCV
 CC proteinase. (Updated on 16-OCT-2003 to standardise OS field)
 XX
 XX Sequence 3010 AA;
 SQ
 Query Match 98.0%; Score 525.5; DB 2; Length 3010;
 Best Local Similarity 99.0%; Pred. No. 3.8e-56;
 Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
QY      1 VPHNIEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKKCDELAALKLTGLGLNAVA 60
DB      1357 VPHNIEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKKCDELAALKLTGLGLNAVA 1416
QY      61 YYRGLDVSVIPTSGD-VVVAATDALMTGFTGDFDSVIDCNCVT 102
DB      1417 YYRGLDVSVIPTSGDVVVVAATDALMTGFTGDFDSVIDCNCVT 1459

RESULT 12
AAR68864
ID AAR68864 standard; protein; 3010 AA.
XX
AC AAR68864;
XX
DT 06-DEC-1995 (first entry)
XX
DE Hepatitis C virus RNA helicase.
XX
KW Hepatitis C virus; HCV; non-A non-B; helicase gene; RNA helicase;
KW baculovirus; recombinant production.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT Region 196..198
FT /label= N-linked glycosylation site
FT Region 209..211
FT /label= N-linked glycosylation site
FT Region 234..236
FT /label= N-linked glycosylation site
FT Region 250..252
FT /label= N-linked glycosylation site
FT Region 305..307
FT /label= N-linked glycosylation site
FT Region 325..327
FT /label= N-linked glycosylation site
FT Region 417..419
FT /label= N-linked glycosylation site
FT Region 423..425
FT /label= N-linked glycosylation site
FT Region 430..432
FT /label= N-linked glycosylation site
FT Region 448..450
FT /label= N-linked glycosylation site
FT Region 532..534
FT /label= N-linked glycosylation site
FT Region 556..558
FT /label= N-linked glycosylation site
FT Region 576..578
FT /label= N-linked glycosylation site
FT Region 623..625
FT /label= N-linked glycosylation site
FT Region 645..647
FT /label= N-linked glycosylation site
FT Region 1213..1215
FT /label= N-linked glycosylation site
FT Region 1255..1257
FT /label= N-linked glycosylation site
FT Region 2041..2043
FT /label= N-linked glycosylation site
FT Region 2077..2079
FT /label= N-linked glycosylation site
FT Region 2240..2242
FT /label= N-linked glycosylation site
FT Region 2788..2790
FT /label= N-linked glycosylation site
XX
XX JP06319583-A.
XX
PD 22-NOV-1994.
XX
```

```
PF 18-SEP-1992; 92JP-00249241.
XX
PR 18-SEP-1992; 92JP-00249241.
XX
PA (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.
XX
DR WPI; 1995-040330/06.
DR N-PSDB; AAQ81559.
XX
of hepatitis C virus helicase gene in baculovirus - useful for large
scale prodn. of RNA helicase.
XX
PS Claim 1; Fig 1-4; 9pp; Japanese.
XX
CC AAQ81559 encodes AAR68864 hepatitis C virus (HCV) RNA helicase. The DNA
CC was used in the construction of an expression vector, which was used to
CC transform a baculovirus host. The transformed baculovirus could then be
CC used for the recombinant prodn. of HCV RNA helicase
XX
SQ Sequence 3010 AA;
Query Match 98.0%; Score 525.5; DB 2; Length 3010;
Best Local Similarity 99.0%; Pred. No. 3.8e-56;
Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 VPHNIEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKKCDELAALKLTGLGLNAVA 60
DB 1357 VPHNIEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKKCDELAALKLTGLGLNAVA 1416
QY 61 YYRGLDVSVIPTSGD-VVVAATDALMTGFTGDFDSVIDCNCVT 102
DB 1417 YYRGLDVSVIPTSGDVVVVAATDALMTGFTGDFDSVIDCNCVT 1459

RESULT 13
AAR82694
ID AAR82694 standard; protein; 3010 AA.
XX
AC AAR82694;
XX
DT 16-OCT-2003 (revised)
DT 14-NOV-1996 (first entry)
XX
DE Partial HCV non-structural polyprotein.
XX
KW proteinase; hepatitis C virus; screening; inhibitor; proteolytic;
KW identification; cleavage.
XX
OS Hepatitis C virus; Virus.
XX
FH Key Location/Qualifiers
FT Protein 898..1233
FT /note= "partial proteinase; see AAR82692"
FT Protein 992..1907
FT /note= "partial proteinase; see AAR82693"
XX
XX JP07184648-A.
XX
PD 25-JUL-1995.
XX
PF 05-FEB-1993; 93JP-00018854.
XX
PR 07-FEB-1992; 92JP-00022657.
PR 18-SEP-1992; 92JP-00249240.
PR 04-DEC-1992; 92JP-00325303.
XX
PA (KAEN/) KAENNO K.
PA (SUMQ) SUMITOMO METAL IND LTD.
PA (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.
XX
DR WPI; 1995-287962/38.
DR N-PSDB; AAT03960.
XX
```

Sequence 3010 AA;

Db 1 VPHNIEEVALSNTGEIPFYGKAIP-EA IKGRHLIFCHSKKCDLAAKLTGLGLNAVA 59

Qy 61 YYRGLDVSVPISGDDVVVATDALTGTGDFDSDVIDCNCVT 102
|||
Db 60 YYRGLDVSVPISGDDVVVATDALTGTGDFDSDVIDCNCVT 101
|||

Search completed: August 12, 2005, 14:14:58
Job time : 109.012 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:04:54 ; Search time 28.57 Seconds
(without alignments)
266.510 Million cell updates/sec

Title: US-09-758-308-2

Perfect score: 536

Sequence: 1 VPHNIEEVALSNTGEIPFY.....LMTGTGDFSDVIDNCVT 102

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/PCITUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	525.5	98.0	278	4	US-08-635-886C-268
2	525.5	98.0	278	4	US-08-974-690C-268
3	514.5	96.0	278	4	US-08-635-886C-272
4	514.5	96.0	278	4	US-08-974-690C-272
5	513.5	95.8	646	3	US-09-198-723A-69
6	513.5	95.8	646	4	US-09-684-881-69
7	512.5	95.6	278	4	US-08-635-886C-273
8	512.5	95.6	278	4	US-08-974-690C-273
9	512.5	95.6	1985	4	US-09-539-601-9
10	512.5	95.6	1985	4	US-09-539-601-12
11	512.5	95.6	1985	4	US-09-539-601-24
12	512.5	95.6	1985	4	US-09-539-601-30
13	512.5	95.6	2201	4	US-09-539-601-6
14	512.5	95.6	2201	4	US-09-539-601-15
15	512.5	95.6	2201	4	US-10-029-907-3
16	512.5	95.6	3010	4	US-09-539-601-3
17	512.5	95.6	3010	4	US-09-539-601-27
18	512.5	95.6	3010	4	US-09-539-601-33
19	508.5	94.9	278	4	US-08-635-886C-271
20	508.5	94.9	278	4	US-08-974-690C-271
21	507.5	94.7	278	4	US-08-635-886C-274
22	507.5	94.7	278	4	US-08-974-690C-274
23	507.5	94.7	609	1	US-08-324-977-40
24	507.5	94.7	609	2	US-08-384-616-40
25	507.5	94.7	609	2	US-08-904-686A-40
26	507.5	94.7	609	3	US-09-315-850-40
27	507.5	94.7	631	1	US-08-700-356-1

28	507.5	94.7	631	2	US-08-936-865-1	Sequence 1, Appli
29	507.5	94.7	632	3	US-09-198-723A-23	Sequence 23, Appl
30	507.5	94.7	632	4	US-09-684-881-23	Sequence 23, Appl
31	507.5	94.7	646	3	US-09-198-723A-60	Sequence 60, Appl
32	507.5	94.7	646	3	US-09-198-723A-63	Sequence 63, Appl
33	507.5	94.7	646	3	US-09-198-723A-66	Sequence 66, Appl
34	507.5	94.7	646	3	US-09-198-723A-72	Sequence 72, Appl
35	507.5	94.7	646	4	US-09-684-881-60	Sequence 60, Appl
36	507.5	94.7	646	4	US-09-684-881-63	Sequence 63, Appl
37	507.5	94.7	646	4	US-09-684-881-66	Sequence 66, Appl
38	507.5	94.7	646	4	US-09-684-881-72	Sequence 72, Appl
39	507.5	94.7	665	4	US-02-543-376B-1	Sequence 1, Appli
40	507.5	94.7	665	4	US-09-543-376B-2	Sequence 2, Appli
41	507.5	94.7	665	4	US-09-543-376B-3	Sequence 3, Appli
42	507.5	94.7	666	3	US-09-198-723A-11	Sequence 11, Appl
43	507.5	94.7	666	3	US-09-198-723A-12	Sequence 12, Appl
44	507.5	94.7	666	3	US-09-198-723A-13	Sequence 13, Appl
45	507.5	94.7	666	3	US-09-198-723A-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-08-635-886C-268
; Sequence 268, Application US/08635886C
; Patent No. 6555114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635.886C
; PRIOR FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 268
; LENGTH: 278
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-268

Query Match 98.0%; Score 525.5; DB 4; Length 278;
Best Local Similarity 99.0%; Pred. NO. 2.5e-58;
Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGEIPFYKAIPIEAKGRHLIFCHSKKCDLAALKLTGLGNVA 60
DB 170 VPHNIEEVALSNTGEIPFYKAIPIEAKGRHLIFCHSKKCDLAALKLTGLGNVA 229
QY 61 YYRGLDVSVIPTSGD-VVWATDALMTGTFDSDVIDNCVT 102
DB 230 YYRGLDVSVIPTSGDVVVWATDALMTGTFDSDVIDNCVT 272

RESULT 2

US-08-974-690C-268
; Sequence 268, Application US/08974690C
; Patent No. 6613333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/08/974.690C


```
; CURRENT FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 268
; LENGTH: 278
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-974-690C-268

Query Match      98.0%; Score 525.5; DB 4; Length 278;
Best Local Similarity 99.0%; Pred. No. 2.5e-58;
Matches 102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHPNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60
DB 170 VPHPNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 229

QY 61 YYRGLDVSIVPTSGD-VVATDALTMTGTFDSDVIDNCNTCVT 102
DB 230 YYRGLDVSIVPTSGD-VVATDALTMTGTFDSDVIDNCNTCVT 272

RESULT 3
US-08-635-886C-272
; Sequence 272, Application US/08635886C
; Patent No. 6555114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635,886C
; CURRENT FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 272
; LENGTH: 278
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-272

Query Match      96.0%; Score 514.5; DB 4; Length 278;
Best Local Similarity 95.1%; Pred. No. 6.2e-57;
Matches 98; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHPNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60
DB 170 VPHPNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 229

QY 61 YYRGLDVSIVPTSGDV-VATDALTMTGTFDSDVIDNCNTCVT 102
DB 230 YYRGLDVSIVPTSGDV-VATDALTMTGTFDSDVIDNCNTCVT 272

RESULT 4
US-08-974-690C-272
; Sequence 272, Application US/08974690C
; Patent No. 6613333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
```

```
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/08/974,690C
; CURRENT FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 272
; LENGTH: 278
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-974-690C-272

Query Match      96.0%; Score 514.5; DB 4; Length 278;
Best Local Similarity 95.1%; Pred. No. 6.2e-57;
Matches 98; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHPNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 60
DB 170 VPHPNIEVALSNTGEIPFYGKAIPLEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVA 229

QY 61 YYRGLDVSIVPTSGDV-VATDALTMTGTFDSDVIDNCNTCVT 102
DB 230 YYRGLDVSIVPTSGDV-VATDALTMTGTFDSDVIDNCNTCVT 272

RESULT 5
US-09-198-723A-69
; Sequence 69, Application US/09198723A
; Patent No. 6211338
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; APPLICANT: Taremi, Shahrar S.
; APPLICANT: Weber, Patricia
; APPLICANT: Yao, Nanhua
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; TITLE OF INVENTION: NS3 Protease and NS4A Cofactor Peptide
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723A
; FILING DATE: 24 NOV 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 646 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-198-723A-69

Query Match      95.8%; Score 513.5; DB 3; Length 646;
```

Best Local Similarity 95.1%; Pred. No. 2.5e-56;
Matches 98; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGEIPFYGKAIPKIEAIGRHLIFCHSKKCKDELAALKLTGLGNAVA 60
DB 346 VPHNIEEVALSNTGEIPFYGKAIPKIEAIGRHLIFCHSKKCKDELAALKLTGLGNAVA 405
QY 61 YRGLDVSIVPTSGD-VVVATDALMTGTFDPSVIDNCVT 102
DB 406 YRGLDVSIVPTSGD-VVVATDALMTGTFDPSVIDNCVT 448

RESULT 6

US-09-684-881-69
Sequence 69, Application US/09684881
Patent No. 6653127

GENERAL INFORMATION:

APPLICANT: Malcolm, Bruce
Taremi, Shahrar S.
Weber, Patricia
Yao, Nanhua

TITLE OF INVENTION: NS3 Protease and NS4A Cofactor Peptide

NUMBER OF SEQUENCES: 123

CORRESPONDENCE ADDRESS:

ADDRESSEE: Schering-Plough Corp.

STREET: 2000 Galloping Hill Road

CITY: Kenilworth

STATE: New Jersey

COUNTRY: USA

ZIP: 07030

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Power Macintosh

OPERATING SYSTEM: 8.0.1

SOFTWARE: Microsoft Word 6.0.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/684,881

FILING DATE: 06-Oct-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/198,723

FILING DATE: 24 Nov 1998

ATTORNEY/AGENT INFORMATION:

NAME: McLaughlin, Jaye P.

REGISTRATION NUMBER: 41,211

REFERENCE/DOCKET NUMBER: JB0800

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908)298-5056

TELEFAX: (908)298-5388

INFORMATION FOR SEQ ID NO: 69:

SEQUENCE CHARACTERISTICS:

LENGTH: 646 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 69:

US-09-684-881-69

Query Match 95.8%; Score 513.5; DB 4; Length 646;
Best Local Similarity 95.1%; Pred. No. 2.5e-56;
Matches 98; Conservative 4; Mismatches 0; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGEIPFYGKAIPKIEAIGRHLIFCHSKKCKDELAALKLTGLGNAVA 60
DB 346 VPHNIEEVALSNTGEIPFYGKAIPKIEAIGRHLIFCHSKKCKDELAALKLTGLGNAVA 405

QY 61 YRGLDVSIVPTSGD-VVVATDALMTGTFDPSVIDNCVT 102
DB 406 YRGLDVSIVPTSGD-VVVATDALMTGTFDPSVIDNCVT 448

RESULT 7

Query Match 95.6%; Score 512.5; DB 4; Length 278;
Best Local Similarity 95.1%; Pred. No. 1.1e-56;
Matches 98; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

US-08-635-886C-273
Sequence 273, Application US/08635886C
Patent No. 6555114

GENERAL INFORMATION:

APPLICANT: LEROUX-ROELS, Geert

APPLICANT: DELSIS, Robert

APPLICANT: MAERTENS, Geert

TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C.

FILE REFERENCE: 2752-18

CURRENT APPLICATION NUMBER: US/08/635,886C

PRIOR FILING DATE: 1996-04-25

PRIOR APPLICATION NUMBER: PCT/EP94/03555

PRIOR FILING DATE: 1994-10-28

PRIOR APPLICATION NUMBER: EP 93402718.6

PRIOR FILING DATE: 1993-11-04

NUMBER OF SEQ ID NOS: 286

SOFTWARE: Patent in version 3.1

SEQ ID NO 273

LENGTH: 278

TYPE: PRT

ORGANISM: hepatitis C virus

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (60)..(60)

OTHER INFORMATION: Xaa is any amino acid

US-08-635-886C-273

Query Match 95.6%; Score 512.5; DB 4; Length 278;
Best Local Similarity 95.1%; Pred. No. 1.1e-56;
Matches 98; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGEIPFYGKAIPKIEAIGRHLIFCHSKKCKDELAALKLTGLGNAVA 60
DB 170 VPHNIEEVALSNTGEIPFYGKAIPKIEAIGRHLIFCHSKKCKDELAALKLTGLGNAVA 229

QY 61 YRGLDVSIVPTSGD-VVVATDALMTGTFDPSVIDNCVT 102
DB 230 YRGLDVSIVPTSGD-VVVATDALMTGTFDPSVIDNCVT 272

RESULT 8

US-08-974-690C-273
Sequence 273, Application US/08974690C
Patent No. 6613333

GENERAL INFORMATION:

APPLICANT: LEROUX-ROELS, Geert

APPLICANT: DELSIS, Robert

APPLICANT: MAERTENS, Geert

TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C

FILE REFERENCE: 2551-94

CURRENT APPLICATION NUMBER: US/08/974,690C

PRIOR FILING DATE: 1997-11-19

PRIOR APPLICATION NUMBER: PCT/EP94/03555

PRIOR FILING DATE: 1994-10-28

PRIOR APPLICATION NUMBER: EP 93402718.6

PRIOR FILING DATE: 1993-11-04

NUMBER OF SEQ ID NOS: 286

SOFTWARE: Patent in version 3.1

SEQ ID NO 273

LENGTH: 278

TYPE: PRT

ORGANISM: hepatitis C virus

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (60)..(60)

OTHER INFORMATION: Xaa is any amino acid

US-08-974-690C-273

Query Match 95.6%; Score 512.5; DB 4; Length 278;
Best Local Similarity 95.1%; Pred. No. 1.1e-56;
Matches 98; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKCDLAALKLTGLGLNAVA 60
Db 170 VPHNIEEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKCDLAALKLTGLGLNAVA 229
QY 61 YYRGLDVSVIPTSGDV-VVATDALMTGTFGDPDSVIDNCNTCVT 102
Db 230 YYRGLDVSVIPTSGDV-VVATDALMTGTFGDPDSVIDNCNTCVT 272

RESULT 9
US-09-539-601-9
; Sequence 9, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; EARLIER FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1985
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-9

Query Match 95.6%; Score 512.5; DB 4; Length 1985;
Best Local Similarity 96.1%; Pred. No. 1.4e-55;
Matches 99; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKCDLAALKLTGLGLNAVA 60
Db 332 VPHNIEEVALSNTGEIPFYGKAIPETIKGRHLIFCHSKKCDLAALKLTGLGLNAVA 391

QY 61 YYRGLDVSVIPTSGDV-VVATDALMTGTFGDPDSVIDNCNTCVT 102
Db 392 YYRGLDVSVIPTSGDV-VVATDALMTGTFGDPDSVIDNCNTCVT 434

RESULT 10
US-09-539-601-12
; Sequence 12, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; EARLIER FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 1985
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-12

Query Match 95.6%; Score 512.5; DB 4; Length 1985;
Best Local Similarity 96.1%; Pred. No. 1.4e-55;
Matches 99; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKCDLAALKLTGLGLNAVA 60
Db 332 VPHNIEEVALSNTGEIPFYGKAIPETIKGRHLIFCHSKKCDLAALKLTGLGLNAVA 391

QY 61 YYRGLDVSVIPTSGDV-VVATDALMTGTFGDPDSVIDNCNTCVT 102

Db 392 YYRGLDVSVIPTSGDV-VVATDALMTGTFGDPDSVIDNCNTCVT 434

RESULT 11
US-09-539-601-24
; Sequence 24, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; EARLIER FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 1985
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-24

Query Match 95.6%; Score 512.5; DB 4; Length 1985;
Best Local Similarity 96.1%; Pred. No. 1.4e-55;
Matches 99; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKCDLAALKLTGLGLNAVA 60
Db 332 VPHNIEEVALSNTGEIPFYGKAIPETIKGRHLIFCHSKKCDLAALKLTGLGLNAVA 391

QY 61 YYRGLDVSVIPTSGDV-VVATDALMTGTFGDPDSVIDNCNTCVT 102
Db 392 YYRGLDVSVIPTSGDV-VVATDALMTGTFGDPDSVIDNCNTCVT 434

RESULT 12
US-09-539-601-30
; Sequence 30, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; EARLIER FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 1985
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-30

Query Match 95.6%; Score 512.5; DB 4; Length 1985;
Best Local Similarity 96.1%; Pred. No. 1.4e-55;
Matches 99; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 VPHNIEEVALSNTGEIPFYGKAIPIEAIKGRHLIFCHSKKCDLAALKLTGLGLNAVA 60
Db 332 VPHNIEEVALSNTGEIPFYGKAIPETIKGRHLIFCHSKKCDLAALKLTGLGLNAVA 391

QY 61 YYRGLDVSVIPTSGDV-VVATDALMTGTFGDPDSVIDNCNTCVT 102
Db 392 YYRGLDVSVIPTSGDV-VVATDALMTGTFGDPDSVIDNCNTCVT 434

RESULT 13
US-09-539-601-6
; Sequence 6, Application US/09539601C
; Patent No. 6630343

```
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2201
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-6

Query Match      95.6%; Score 512.5; DB 4; Length 2201;
Best Local Similarity 96.1%; Pred. No. 1.6e-55;
Matches 99; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY      1 VPHPNIEEVALSNTGGEIPFYGKAIPTEAIKGRHILFCHSKKCDLAALKLTGLGLNAVA 60
DB      548 VPHPNIEEVALSNTGGEIPFYGKAIPTEAIKGRHILFCHSKKCDLAALKLTGLGLNAVA 607

QY      61 YRGLDVSIVPTSGDV-VVATDALMTGFTGDFSDVIDNCNVCVT 102
DB      608 YRGLDVSIVPTSGDV-VVATDALMTGFTGDFSDVIDNCNVCVT 650

RESULT 14
US-09-539-601-15
; Sequence 15, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 2201
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-15

Query Match      95.6%; Score 512.5; DB 4; Length 2201;
Best Local Similarity 96.1%; Pred. No. 1.6e-55;
Matches 99; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY      1 VPHPNIEEVALSNTGGEIPFYGKAIPTEAIKGRHILFCHSKKCDLAALKLTGLGLNAVA 60
DB      548 VPHPNIEEVALSNTGGEIPFYGKAIPTEAIKGRHILFCHSKKCDLAALKLTGLGLNAVA 607

QY      61 YRGLDVSIVPTSGDV-VVATDALMTGFTGDFSDVIDNCNVCVT 102
DB      608 YRGLDVSIVPTSGDV-VVATDALMTGFTGDFSDVIDNCNVCVT 650

RESULT 15
US-10-029-907-3
; Sequence 3, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; FILE REFERENCE: HEPATITIS C VIRUS
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
```

```
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2201
; TYPE: PRT
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 882
; OTHER INFORMATION: Xaa is Lys or Arg
; NAME/KEY: VARIANT
; LOCATION: 1489
; OTHER INFORMATION: Xaa is Leu
US-10-029-907-3
```

```
Query Match      95.6%; Score 512.5; DB 4; Length 2201;
Best Local Similarity 96.1%; Pred. No. 1.6e-55;
Matches 99; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY      1 VPHPNIEEVALSNTGGEIPFYGKAIPTEAIKGRHILFCHSKKCDLAALKLTGLGLNAVA 60
DB      548 VPHPNIEEVALSNTGGEIPFYGKAIPTEAIKGRHILFCHSKKCDLAALKLTGLGLNAVA 607

QY      61 YRGLDVSIVPTSGDV-VVATDALMTGFTGDFSDVIDNCNVCVT 102
DB      608 YRGLDVSIVPTSGDV-VVATDALMTGFTGDFSDVIDNCNVCVT 650
```

Search completed: August 12, 2005, 14:24:48
Job time : 29.57 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:00:18 ; Search time 95.4717 Seconds
(without alignments)
368.645 Million cell updates/sec

Title: US-09-758-308-1

Perfect score: 501

Sequence: 1 MSTNPKPQKTKRNTNRRPQ.....EGRTWAQGYWPPLYGNEG 91

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980a:*
- 2: geneseqp1990a:*
- 3: geneseqp2000a:*
- 4: geneseqp2001a:*
- 5: geneseqp2002a:*
- 6: geneseqp2003a:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	501	100.0	91	4	AAB31695 Antigenic
2	501	100.0	124	3	AAY82358 Hepatitis
3	501	100.0	161	2	Aaw33689 Recombina
4	501	100.0	191	2	Aar24525 HCV in pl
5	501	100.0	191	2	Aar92954 Hepatitis
6	501	100.0	191	4	Aab73170 Protein #
7	501	100.0	263	2	Aar24527 HCV in pl
8	501	100.0	281	2	Aaw33691 Nucleic a
9	501	100.0	980	2	Aar53921 HCV fusio
10	501	100.0	1010	2	Aar88311 HCV pepi
11	501	100.0	3010	2	Aar68864 Hepatitis
12	501	100.0	3010	2	Aar82694 Partial H
13	501	100.0	3010	7	Adf88597 Hepatitis
14	498	99.4	97	2	Aar40978 HCV core
15	498	99.4	154	2	Aaw30583 Hepatitis
16	498	99.4	154	2	Aay01614 Protein e
17	498	99.4	154	2	Adf66059 Hepatitis
18	498	99.4	191	2	Aar35758 Core prot
19	498	99.4	191	2	Aar56597 HCV core
20	498	99.4	223	2	Aar54641 HCV EN-80
21	498	99.4	223	2	Aaw18199 Hepatitis
22	498	99.4	223	2	Aaw07484 HCV unpro
23	497	99.2	120	2	Aaw33687 Hepatitis
24	497	99.2	150	2	Aaw33688 Hepatitis
25	497	99.2	176	2	Aaw33695 Mouse pro

26	497	99.2	191	2	AAR92955	Aar92955 Hepatitis
27	497	99.2	191	2	AAR92953	Aar92953 Hepatitis
28	497	99.2	191	2	Aaw33690	Aaw33690 Recombina
29	497	99.2	326	2	Aaw68461	Aaw68461 Protein e
30	497	99.2	423	2	Aaw68465	Aaw68465 Protein e
31	497	99.2	845	2	Aaw68466	Aaw68466 Protein e
32	496	99.0	103	5	ABB77253	Abb77253 HCV bait
33	496	99.0	105	3	AB18537	Ab18537 Protein e
34	496	99.0	113	5	ABB77254	Abb77254 HCV bait
35	496	99.0	116	8	ADN35970	Adn35970 HCV CDNA
36	496	99.0	120	8	ADL16316	Adl16316 Hepatitis
37	496	99.0	151	7	ADP08480	Adp08480 Hepatitis
38	496	99.0	154	2	AAY14970	Aay14970 HCV J7 C/
39	496	99.0	173	7	ADN33109	Adn33109 Hepatitis
40	496	99.0	178	2	AAR92947	Aar92947 Hepatitis
41	496	99.0	182	5	AAE19889	Aae19889 Hepatitis
42	496	99.0	182	7	ABW00340	Abw00340 Hepatitis
43	496	99.0	191	2	AAR44010	Aar44010 Hepatitis
44	496	99.0	191	2	AAW41754	Aaw41754 Hepatitis
45	496	99.0	191	2	AAR92938	Aar92938 Hepatitis

ALIGNMENTS

RESULT 1

AAB31695
ID AAB31695 standard; peptide; 91 AA.

AC AAB31695;

DT 30-APR-2001 (first entry)

DE Antigenic epitope of the Hepatitis C virus (HCV) core protein.

XX Antigen; HCV; polyprotein; core protein; NS3 protein; NS4 protein;

KW NS4a protein; HCV infection.

OS Hepatitis C virus.

PN WO200104149-A1.

PD 18-JAN-2001.

PF 07-JUL-2000; 2000WO-US018704.

PR 09-JUL-1999; 99WO-US015578.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Fields HA, Khudyakov YE;

DR WPI; 2001-138316/14.

XX New (mosaic) polypeptides, useful as reagents in assays for the diagnosis or monitoring of HCV infections and as components of anti-HCV vaccines, comprises antigenic groups of hepatitis C virus (HCV).

PS Claim 2; Page 39; 52pp; English.

XX The present sequence represents an antigenic epitope from a hepatitis C virus (HCV) core protein. The peptide comprises amino acids 1-91 of the HCV polyprotein. The specification describes antigenic epitopes from HCV core protein, NS3 protein, NS4 protein and NS4a protein. The antigenic peptides are useful as diagnostic reagents for detecting HCV in a biological sample. They are also useful for monitoring HCV infection in a patient sample in addition to diagnosis. Pharmaceutical compositions comprising the peptides are useful for preventing, minimizing or reducing HCV infection in patients who have been exposed to HCV or to individuals, such as health care workers or blood product recipients, who are more likely to become exposed to HCV infection. The antigenic peptides are also useful for generating antibodies which can be used to detect HCV proteins in a sample or for laboratory research purposes

```
XX SQ Sequence 91 AA;
Query Match 100.0%; Score 501; DB 4; Length 91;
Best Local Similarity 100.0%; Pred. No. 8.9e-47;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSTNPKPQKTKRNTNRPPQDVKPGGGQIVGGVYLLPRRGLGVRAATKTSRSQPRG 60
DB 1 MSTNPKPQKTKRNTNRPPQDVKPGGGQIVGGVYLLPRRGLGVRAATKTSRSQPRG 60

OY 61 RROPIPKARRPEGRRTWAQPGYPWPPLYGNEG 91
DB 61 RROPIPKARRPEGRRTWAQPGYPWPPLYGNEG 91

RESULT 3
AAW33689
ID AAW33689 standard; protein; 161 AA.
XX
AC AAW33689;
XX
DT 30-APR-1998 (first entry)
XX
DE Recombinant protein 120NA.
XX
Nucleic acid-binding motif; HCV; HBV; Hepatitis C virus; diagnosis;
KW core polypeptide; immunoassay; detection; antigen; disease; infection;
XX Hepatitis B virus; recombinant.
OS Synthetic.
OS Hepatitis C virus.
OS Hepatitis B virus.
OS Chimeric.
XX
PN EP805160-A1.
XX
PD 05-NOV-1997.
XX
PF 30-APR-1997; 97EP-00400985.
XX
PR 01-MAY-1996; 96JP-00134444.
XX
PA (FURE ) FUJIREBIO INC.
XX
PI Takemura F, Ueno E, Itoh S;
XX
WPI; 1997-529030/49.
DR N-PSDB; AAW06337.
XX
PT Nucleic acid-bound polypeptide - useful as immunoassay reagent.
XX
PS Example 1; Page 20-21; 38pp; English.
XX
CC This is a recombinant protein 120NA. This recombinant protein contains
CC the Hepatitis C virus (HCV) core polypeptide 120 fused to a nucleic acid-
CC binding motif present in the HBC protein of Hepatitis B virus (HBV). This
CC recombinant protein can be bound to a nucleic acid in the host for
CC producing a nucleic acid-bound polypeptide by a new method. The method
CC comprises producing a fusion gene containing the polypeptide and the
CC nucleic acid-binding motif, binding a nucleic acid to the polypeptide as
CC a soluble fraction, and purifying the nucleic acid-bound polypeptide from
CC the soluble fraction. When the polypeptide is a recombinant form of an
CC antigen, the nucleic acid-bound polypeptide can be used as an immunoassay
CC reagent for detecting the antigen or an antibody to the antigen,
CC especially in an agglutination assay using particles coated with the
CC nucleic acid-bound polypeptide. The methods can be applied to diagnosis
CC of disease and infection, especially for the detection of HBV and HCV
CC polypeptides. The nucleic acid-bound polypeptides may be immunoreactive
CC in cases where the free polypeptide is not
XX
SQ Sequence 161 AA;
Query Match 100.0%; Score 501; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.7e-46;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSTNPKPQKTKRNTNRPPQDVKPGGGQIVGGVYLLPRRGLGVRAATKTSRSQPRG 60
DB 6 MSTNPKPQKTKRNTNRPPQDVKPGGGQIVGGVYLLPRRGLGVRAATKTSRSQPRG 65

OY 61 RROPIPKARRPEGRRTWAQPGYPWPPLYGNEG 91
DB 61 RROPIPKARRPEGRRTWAQPGYPWPPLYGNEG 91

6
```


Db 66 RQPIPKARRPEGRTWAQPGYPWPLYGNEG 96

RESULT 4

ID AAR24525 standard; protein; 191 AA.

XX AAR24525;

XX 07-DEC-1992 (first entry)

XX HCV in plasmid pKMR3.

XX HCV; antibody; agglutination; ELISA.

XX Hepatitis C virus.

XX JP04144686-A.

XX 19-MAY-1992.

XX 04-OCT-1990; 90JP-00265228.

XX 04-OCT-1990; 90JP-00265228.

XX (TOKU) TOKUYAMA SODA KK.

XX (SHIMU) SHIMOTONO K.

XX WPI; 1992-214123/26.

XX N-PSDB; AAQ25609.

XX New structural protein gene of hepatitis C virus - encodes specific sequence of 163 aminoacid(s), useful for the diagnosis of hepatitis C.

XX Disclosure; Fig 2; 13pp; Japanese.

XX The sequences given in AAR24525-27 are examples of a new structural gene of hepatitis C virus (HCV) encoded by various expression vectors. This polypeptide has HCV antigen activity and can be used to recognise anti-HCV antibodies existing in the serum of HCV patients specifically. It can be used as a diagnostic agent by the agglutination or ELISA method

XX Sequence 191 AA;

Query Match 100.0%; Score 501; DB 2; Length 191;

Best Local Similarity 100.0%; Pred. No. 2e-46; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKNTNRRPQDVKFGGGQIVGGVLLPRGPRGLGVRATKTSERSQPRG 60

Db 1 MSTNPKPQKTKNTNRRPQDVKFGGGQIVGGVLLPRGPRGLGVRATKTSERSQPRG 60

QY 61 RQPIPKARRPEGRTWAQPGYPWPLYGNEG 91

Db 61 RQPIPKARRPEGRTWAQPGYPWPLYGNEG 91

RESULT 5

ID AAR92954 standard; protein; 191 AA.

XX AAR92954;

XX 02-OCT-1996 (first entry)

XX Hepatitis C virus isolate HK5 core protein.

XX HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine; hepatitis.

XX Hepatitis C virus.

XX WO9605315-A2.

XX 22-FEB-1996.

XX 15-AUG-1995; 95WO-US010398.

XX 15-AUG-1994; 94US-00290665.

XX (USSH) US SEC DEPT HEALTH.

XX Bukh J, Miller RH, Purcell RH;

XX WPI; 1996-139709/14.

XX N-PSDB; AAT16628.

XX DNA and amino acid sequence of HCV envelope 1 and core proteins - used to determine HCV genotype and as vaccines against HCV infection.

XX Claim 4; Page 195-196; 340pp; English.

XX AAR92936-R32987 are HCV core proteins derived from 52 different HCV isolates. Isolated cDNA sequences are used for the prodn. of primers useful for detecting the presence of HCV in a sample, the primers are also useful for HCV genotyping. Proteins encoded by the cDNAs can be used in vaccines for immunising against HCV infection. The proteins may also be used to detect antibodies against HCV in serum, saliva, lymphocytes or other mononuclear cells. The antibodies may be used in the prevention of HCV infection

XX Sequence 191 AA;

Query Match 100.0%; Score 501; DB 2; Length 191;

Best Local Similarity 100.0%; Pred. No. 2e-46;

Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKNTNRRPQDVKFGGGQIVGGVLLPRGPRGLGVRATKTSERSQPRG 60

Db 1 MSTNPKPQKTKNTNRRPQDVKFGGGQIVGGVLLPRGPRGLGVRATKTSERSQPRG 60

QY 61 RQPIPKARRPEGRTWAQPGYPWPLYGNEG 91

Db 61 RQPIPKARRPEGRTWAQPGYPWPLYGNEG 91

RESULT 6

AAB73170

ID AAB73170 standard; protein; 191 AA.

XX AAB73170;

XX 06-AUG-2003 (revised)

DT 11-MAY-2001 (first entry)

XX Protein #1: SEQ ID 2.

XX Viral infection; RecQ helicase; liver disorder; HCV infection; liver cancer.

XX Hepatitis C virus.

XX WO200111089-A1.

XX 15-FEB-2001.

XX 03-AUG-2000; 2000WO-JP005218.

XX 05-AUG-1999; 99JP-00223097.

XX (AGEN-) AGENE RES INST CO LTD.

XX Furuichi Y, Shimamoto A;

XX WPI; 2001-191559/19.

XX N-PSDB; AAF63461.

XX This sequence is encoded by a fragment of the hepatitis C virus (HCV)
 CC genome including the hypervariable region. The DNA region is amplified by
 CC the primer sequences given in AAR03664-73. The amplified product is
 CC subjected to electrophoresis under denaturing conditions. The primers may
 CC be used in combination with the primer sequences given in AAR03674-76.
 CC Preferably, primer MS1, MS2, MS3, MS4, MS5 or MS6 and an oligo selected
 CC from MR1, MR2 or MR1' are used as primer pairs
 XX

SQ Sequence 1010 AA;

Query Match 100.0%; Score 501; DB 2; Length 1010;
 Best Local Similarity 100.0%; Pred. No. 1.2e-45;
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQRTKRNTRRPQDVKFGGQIVGGVYLLPRRGLGVRATRTKTSRSQPRG 60
 |||||
 DB 1 MSTNPKPQRTKRNTRRPQDVKFGGQIVGGVYLLPRRGLGVRATRTKTSRSQPRG 60
 |||||

QY 61 RROPIPKARRPEGRRTWAQPGYPWPLYGNEGM 91
 |||||

DB 61 RROPIPKARRPEGRRTWAQPGYPWPLYGNEGM 91
 |||||

RESULT 11
 AAR68864
 ID AAR68864 standard; protein; 3010 AA.

XX AC AAR68864;

XX DT 06-DEC-1995 (first entry)

XX DE Hepatitis C virus RNA helicase.

XX KW Hepatitis C virus; HCV; non-A non-B; helicase gene; RNA helicase;
 KW baculovirus; recombinant production.

XX OS Hepatitis C virus.

XX FH Key Location/Qualifiers
 FT Region 196..198
 FT /label= N-linked glycosylation site
 FT Region 209..211
 FT /label= N-linked glycosylation site
 FT Region 234..236
 FT /label= N-linked glycosylation site
 FT Region 250..252
 FT /label= N-linked glycosylation site
 FT Region 305..307
 FT /label= N-linked glycosylation site
 FT Region 325..327
 FT /label= N-linked glycosylation site
 FT Region 417..419
 FT /label= N-linked glycosylation site
 FT Region 423..425
 FT /label= N-linked glycosylation site
 FT Region 430..432
 FT /label= N-linked glycosylation site
 FT Region 448..450
 FT /label= N-linked glycosylation site
 FT Region 532..534
 FT /label= N-linked glycosylation site
 FT Region 556..558
 FT /label= N-linked glycosylation site
 FT Region 576..578
 FT /label= N-linked glycosylation site
 FT Region 623..625
 FT /label= N-linked glycosylation site
 FT Region 645..647
 FT /label= N-linked glycosylation site
 FT Region 1213..1215
 FT /label= N-linked glycosylation site
 FT Region 1255..1257

FT Region /label= N-linked glycosylation site
 FT 2041..2043
 FT /label= N-linked glycosylation site
 FT Region 2077..2079
 FT /label= N-linked glycosylation site
 FT Region 2240..2242
 FT /label= N-linked glycosylation site
 FT Region 2788..2790
 FT /label= N-linked glycosylation site
 XX
 XX JP06319583-A.

XX PD 22-NOV-1994.

XX PF 18-SEP-1992; 92JP-00249241.

XX PR 18-SEP-1992; 92JP-00249241.

XX PA (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.

XX DR WPI; 1995-040330/06.

XX DR N-PSDB; AAQ81559.

XX of hepatitis C virus helicase gene in baculovirus - useful for large
 PT scale prodn. of RNA helicase.

XX PS Claim 1; Fig 1-4; 9pp; Japanese.

XX CC AAQ81559 encodes AAR68864 hepatitis C virus (HCV) RNA helicase. The DNA
 CC was used in the construction of an expression vector, which was used to
 CC transform a baculovirus host. The transformed baculovirus could then be
 CC used for the recombinant prodn. of HCV RNA helicase
 XX

SQ Sequence 3010 AA;

Query Match 100.0%; Score 501; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 3.8e-45;
 Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQRTKRNTRRPQDVKFGGQIVGGVYLLPRRGLGVRATRTKTSRSQPRG 60
 |||||
 DB 1 MSTNPKPQRTKRNTRRPQDVKFGGQIVGGVYLLPRRGLGVRATRTKTSRSQPRG 60
 |||||

QY 61 RROPIPKARRPEGRRTWAQPGYPWPLYGNEGM 91
 |||||

DB 61 RROPIPKARRPEGRRTWAQPGYPWPLYGNEGM 91
 |||||

RESULT 12

AAR82694
 ID AAR82694 standard; protein; 3010 AA.

XX AC AAR82694;

XX DT 16-OCT-2003 (revised)

XX DT 14-NOV-1996 (first entry)

XX DE Partial HCV non-structural polypeptide.

XX KW proteinase; hepatitis C virus; screening; inhibitor; proteolytic;
 KW identification; cleavage.

XX OS Hepatitis C virus; Virus.

XX FH Key Location/Qualifiers

FT Protein 898..1233

FT /note= "partial proteinase; see AAR82692"

FT Protein 992..1907

FT /note= "partial proteinase; see AAR82693"

XX JP07184648-A.

XX 25-JUL-1995.

XX 05-FEB-1993; 93JP-00018854.
XX
XX 07-FEB-1992; 92JP-00022657.
PR 18-SEP-1992; 92JP-00249240.
PR 04-DEC-1992; 92JP-00325303.
XX
XX (KAEN/) KAENNO K.
PA (SUMQ) SUMITOMO METAL IND LTD.
PA (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.
XX
XX WPI; 1995-287962/38.
DR N-PSDB; AAT03960.
XX
XX An HCV proteinase active substance - which has activity as an anti-HCV
PT agent and can be used to screen for proteinase inhibitors.
XX
XX Disclosure; Page 39-48; 52pp; Japanese.
XX
XX The present sequence is a partial Hepatitis C Virus (HCV) polyprotein
CC from the non-structural region. Partial proteinase sequences (AAR82692-
CC 93) are contained within this sequence. The proteinases can be used as
CC anti-HCV agents. They can also be used to screen cpds. for their ability
CC to inhibit their proteolytic activity. In this way proteinase inhibitors
CC can be identified. (Updated on 16-OCT-2003 to standardise OS field)
XX
XX Sequence 3010 AA;
SQ

Query Match 100.0%; Score 501; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 3.8e-45;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKPGGQIVGGVLLPRGPRGLGVTRTKTSERSQPRG 60
DB 1 MSTNPKPQKTKRNTNRRPQDVKPGGQIVGGVLLPRGPRGLGVTRTKTSERSQPRG 60
QY 61 RROQIPKARRPEGRTWAQPGYPWPLYGNEGM 91
DB 61 RROQIPKARRPEGRTWAQPGYPWPLYGNEGM 91

RESULT 13
ADP88597
ID ADP88597 standard; protein; 3010 AA.
XX
XX ADP88597;
AC
DT 26-FEB-2004 (first entry)
XX
XX Hepatitis C virus NS3 gene protein, SEQ ID NO 6.
XX liver cancer; hepatitis-C virus; NS3 gene; carcinogenicity.
XX Hepatitis C virus.
XX JP2003210181-A.
XX
XX 29-JUL-2003.
XX
XX 30-MAY-2002; 2002JP-00158335.
XX
XX 16-NOV-2001; 2001JP-00352443.
XX
XX (SHIN-) ZH SHINSANGYO SOZO KENKYU KIKO.
XX
XX WPI; 2003-819836/77.
DR N-PSDB; ADP88596.
XX
XX Diagnosing liver cancer, involves amplifying amino terminal region of
PT hepatitis-C virus gene using predetermined primer and determining
PT hepatitis-C virus in base sequence of obtained DNA fragment.
XX
XX Disclosure; SEQ ID NO 6; 36pp; Japanese.
PS

XX The invention relates to the novel testing method for diagnosing liver
CC cancer. The novel method comprises amplifying the amino terminal region
CC of a hepatitis-C virus NS3 gene using a predetermined primer and
CC determining the hepatitis-C virus in a base sequence of the obtained DNA
CC fragment. The novel testing method is useful for diagnosing liver cancer
CC and also used in a gene amplification technique, a clinical laboratory
CC test reagent, a polymerase chain reaction, a base sequence analysis and
CC genetic engineering. The method enables the detection of a hepatitis-C
CC virus having high carcinogenicity with high specificity. This sequence
CC represents the protein of the hepatitis-C virus NS3 gene of the
CC invention.
XX
XX Sequence 3010 AA;
SQ

Query Match 100.0%; Score 501; DB 7; Length 3010;
Best Local Similarity 100.0%; Pred. No. 3.8e-45;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKPGGQIVGGVLLPRGPRGLGVTRTKTSERSQPRG 60
DB 1 MSTNPKPQKTKRNTNRRPQDVKPGGQIVGGVLLPRGPRGLGVTRTKTSERSQPRG 60
QY 61 RROQIPKARRPEGRTWAQPGYPWPLYGNEGM 91
DB 61 RROQIPKARRPEGRTWAQPGYPWPLYGNEGM 91

RESULT 14
AAR40978
ID AAR40978 standard; peptide; 97 AA.
XX
XX AAR40978;
AC
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 02-MAR-1994 (first entry)
XX
XX HCV core protein N-terminal fragment.
XX
XX Hepatitis C virus; HCV; epitope; core region; detection; vaccine.
XX Hepatitis C virus.
XX WO9317111-A1.
XX
XX 02-SEP-1993.
XX
XX 26-FEB-1993; 93WO-GB000410.
XX
XX 28-FEB-1992; 92GB-00004274.
XX
XX (WELL) WELLCOME FOUND LTD.
XX
XX Pike I;
XX
XX WPI; 1993-288416/36.
XX
XX New hepatitis C virus peptide(s) and antibodies - used for immunoassays
PT for hepatitis C virus or protective vaccines.
XX
XX Example 1; Page 27; 43pp; English.
XX
XX This sequence represents the N-terminal 97 amino acids of the hepatitis C
CC virus (HCV) core protein. Peptide fragments derived from this sequence
CC contained epitopic regions. These peptides may be used alone or as a part
CC of a larger peptide, or in a mixture of peptides for the detection of HCV
CC or in a vaccine for its prevention. See also AAR40979-99. (Updated on 25-
CC MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
CC field.)
XX
XX Sequence 97 AA;
SQ

Query Match 99.4%; Score 498; DB 2; Length 97;
 Best Local Similarity 98.9%; Pred. No. 2e-46;
 Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRTKTSERSQPRG 60
 |||||
 Db 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRTKTSERSQPRG 60
 |||||

QY 61 RRQPIPKARRPEGRRTWAQPGYPWPPLYGNEGM 91
 |||||
 Db 61 RRQPIPKARRPEGRRTWAQPGYPWPPLYGNEGL 91
 |||||

RESULT 15

AAW30583
 ID AAW30583 standard; protein; 154 AA.

XX AC AAW30583;

XX DT 10-MAY-1999 (first entry)

XX DE Hepatitis C virus J7 isolate C/E consensus sequence.

XX KW HCV; non-A, non-B hepatitis; NANBH; diagnosis; vaccine; antigen;
 antibody; immunoassay; assay; core protein; envelope protein.

XX OS Hepatitis C virus.

XX PH Key Location/Qualifiers
 FT Misc-difference 8 /note= "Glu in clone 1"
 FT Misc-difference 25 /note= "Leu in clone 1"
 FT Misc-difference 91 /note= "END in clone b"
 FT Misc-difference 110 /note= "Thr in clone b"

XX PN US5856437-A.

XX PD 05-JAN-1999.

XX PF 03-NOV-1994; 94US-00334255.

XX PR 15-SEP-1989; 89US-00408045.

XX PR 21-DEC-1989; 89US-00456142.

XX PR 04-JAN-1991; 91US-00637380.

XX PR 02-AUG-1993; 93US-00101280.

XX PR 24-FEB-1994; 94US-00201066.

XX PA (NAHE-) NAT INST OF HEALTH JAPAN.

XX PI (CHIR) CHIRON CORP.

XX PI Weiner AJ, Miyamura T, Irvine BD, Kolberg JA, Han J, Saito I;
 Cha T, Houghton M;

XX DR WPI; 1999-105191/09.

XX DR N-PSDB; AAW00400.

XX PT Antigenic polypeptides from J1 and J7 hepatitis C virus isolates - useful
 as immunoassay reagents, for raising antibodies and as vaccine
 components.

XX PS Claim 1; Fig 6; 44pp; English.

XX CC This is a consensus sequence of a fragment of the core/envelope (C/E)
 domain of hepatitis C virus (HCV) isolate J7. New isolates J1 and J7 of
 HCV were isolated from Japanese blood donors who had been implicated as
 non-A, non-B hepatitis (NANBH) carriers, and the amino acid sequences of
 various viral domains of these isolates were determined (see AAW30583-
 87). The C/E region of J1 has 97.4% amino acid homology to the HCV1
 prototype C/E region. Heterogeneity was detected between J7 C/E clones
 isolated from the same donor. The new HCV isolates have amino acid and

CC nucleotide (see also AAX00400-04) sequences which are distinct from the
 CC HCV1 prototype. These differences can be exploited for use in diagnostics
 CC for NANBH, recombinant protein production and vaccine development.
 CC Claimed HCV J1 and J2 polypeptide antigens can be used: (i) as
 CC immunoassay reagents, or standards, to detect HCV antibodies, e.g. for
 CC diagnosing infection or screening donated blood; (ii) to generate
 CC specific antibodies (used for detecting the corresponding polypeptide, to
 CC screen for antiviral agents, for virus isolation and for passive
 CC immunisation); (iii) in protective or therapeutic vaccines, and (iv) for
 CC isolation of non-A, non-B viruses
 XX
 SQ Sequence 154 AA;

Query Match 99.4%; Score 498; DB 2; Length 154;
 Best Local Similarity 98.9%; Pred. No. 3.3e-46;

Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRTKTSERSQPRG 60
 |||||

Db 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRTKTSERSQPRG 60
 |||||

QY 61 RRQPIPKARRPEGRRTWAQPGYPWPPLYGNEGM 91
 |||||

Db 61 RRQPIPKARRPEGRRTWAQPGYPWPPLYGNEGL 91
 |||||

Search completed: August 12, 2005, 14:14:56

Job time : 97.4717 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:04:54 ; Search time 25.4889 Seconds
(without alignments)
266.510 Million cell updates/sec

Title: US-09-758-308-1
Perfect score: 501
Sequence: 1 MSTNPKPQKTKRNTRRPQ.....EGRTWAQPGYWPVLYGNEGM 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	501	100.0	191	2	US-08-290-665A-173
2	501	100.0	191	5	PCT-US95-10398-173
3	501	100.0	450	4	US-08-635-886C-187
4	501	100.0	450	4	US-08-635-886C-187
5	498	99.4	190	1	US-07-681-701-16
6	498	99.4	223	1	US-08-143-579A-4
7	498	99.4	223	3	US-08-143-578A-4
8	498	99.4	223	3	US-08-454-928-8
9	498	99.4	340	1	US-08-462-195-2
10	498	99.4	340	2	US-08-636-883-2
11	498	99.4	340	3	US-09-127-823-2
12	497	99.2	191	2	US-08-290-665A-172
13	497	99.2	191	2	US-08-290-665A-174
14	497	99.2	191	5	PCT-US95-10398-172
15	497	99.2	191	5	PCT-US95-10398-174
16	496	99.0	120	4	US-08-931-855B-10
17	496	99.0	182	4	US-10-104-966-2
18	496	99.0	191	2	US-08-290-665A-156
19	496	99.0	191	2	US-08-290-665A-157
20	496	99.0	191	2	US-08-290-665A-158
21	496	99.0	191	2	US-08-290-665A-159
22	496	99.0	191	2	US-08-290-665A-160
23	496	99.0	191	2	US-08-290-665A-164
24	496	99.0	191	2	US-08-290-665A-166
25	496	99.0	191	3	US-08-380-160-3
26	496	99.0	191	5	PCT-US95-10398-156
27	496	99.0	191	5	PCT-US95-10398-157

Sequence 158, App
Sequence 159, App
Sequence 160, App
Sequence 164, App
Sequence 166, App
Sequence 179, App
Sequence 180, App
Sequence 180, App
Sequence 179, App
Sequence 180, App
Sequence 190, App
Sequence 50, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 12, Appl
Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-290-665A-173
; Sequence 173, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 173:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: HK5
; US-08-290-665A-173
Query Match 100.0%; Score 501; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 2.6e-51;

Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPQKTKRNTNRRPQDVKPGGQIVGGVYLLPRRGLGVTRATKTSRSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKPGGQIVGGVYLLPRRGLGVTRATKTSRSQPRG 60

Qy 61 RROPIPKARRPEGRRTWAQPGYPWPPLYGNEG 91
Db 61 RROPIPKARRPEGRRTWAQPGYPWPPLYGNEG 91

RESULT 2
PCT-US95-10398-173
; Sequence 173, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10398
; FILING DATE: 15-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086.428
; FILING DATE: 29 JUNE 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290/665
; FILING DATE: 15 AUGUST 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 173:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: HK5
PCT-US95-10398-173

Query Match 100.0%; Score 501; DB 5; Length 191;
Best Local Similarity 100.0%; Pred. No. 2.6e-51;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPQKTKRNTNRRPQDVKPGGQIVGGVYLLPRRGLGVTRATKTSRSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKPGGQIVGGVYLLPRRGLGVTRATKTSRSQPRG 60

Qy 61 RROPIPKARRPEGRRTWAQPGYPWPPLYGNEG 91

Db 61 RROPIPKARRPEGRRTWAQPGYPWPPLYGNEG 91

RESULT 3
US-08-635-886C-187
; Sequence 187, Application US/08635886C
; Patent No. 555114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635.886C
; CURRENT FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 187
; LENGTH: 450
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-187

Query Match 100.0%; Score 501; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. No. 7.3e-51;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPQKTKRNTNRRPQDVKPGGQIVGGVYLLPRRGLGVTRATKTSRSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKPGGQIVGGVYLLPRRGLGVTRATKTSRSQPRG 60

Qy 61 RROPIPKARRPEGRRTWAQPGYPWPPLYGNEG 91
Db 61 RROPIPKARRPEGRRTWAQPGYPWPPLYGNEG 91

RESULT 4
US-08-974-690C-187
; Sequence 187, Application US/08974690C
; Patent No. 661333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/08/974.690C
; CURRENT FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 187
; LENGTH: 450
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-974-690C-187

Query Match 100.0%; Score 501; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. No. 7.3e-51;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPQKTKRNTNRRPQDVKPGGQIVGGVYLLPRRGLGVTRATKTSRSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKPGGQIVGGVYLLPRRGLGVTRATKTSRSQPRG 60

Db 1 MSTNPKPQKTKRNTNRRPDQVKFPGGGQIVGGVLLPRRGPRLGVRATRKTSERSQPRG 60

QY 61 RROQIPKARRPGRRTWAQPGYWPPLYGNEG 91

Db 61 RROQIPKARRPGRRTWAQPGYWPPLYGNEG 91

RESULT 5

US-07-681-701-16

Sequence 16, Application US/07681701

Patent No. 5574132

GENERAL INFORMATION:

APPLICANT: Lacroix, Martial

TITLE OF INVENTION: PEPTIDES AND MIXTURES THEREOF FOR

TITLE OF INVENTION:

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: FISH & NEAVE

STREET: 875 Third Avenue

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/681,701

FILING DATE: 19910405

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Haley Jr., James F.

REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: IAF-10

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 715-0742

TELEFAX: (212) 715-0673

TELEX: 14-8367

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 190 amino acids

TYPE: AMINO ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-07-681-701-16

Query Match 99.4%; Score 498; DB 1; Length 190;

Best Local Similarity 98.9%; Pred. No. 5.8e-51;

Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPDQVKFPGGGQIVGGVLLPRRGPRLGVRATRKTSERSQPRG 60

Db 1 MSTNPKPQKTKRNTNRRPDQVKFPGGGQIVGGVLLPRRGPRLGVRATRKTSERSQPRG 60

QY 61 RROQIPKARRPGRRTWAQPGYWPPLYGNEG 91

Db 61 RROQIPKARRPGRRTWAQPGYWPPLYGNEG 91

RESULT 6

US-08-143-579A-4

Sequence 4, Application US/08143579A

Patent No. 5625034

GENERAL INFORMATION:

APPLICANT: Liao, Jaw-Ching

TITLE OF INVENTION: CORE ANTIGEN PROTEIN OF HEPATITIS C

TITLE OF INVENTION: VIRUS, AND DIAGNOSTIC METHOD AND KIT USING THE SAME

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: U.S.

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/143,578A

FILING DATE: 26-OCT-1993

CLASSIFICATION: 435

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: U.S.

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/143,579A

FILING DATE: 26-OCT-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/963483

FILING DATE: 16-OCT-1992

ATTORNEY/AGENT INFORMATION:

NAME: King, Joshua

REGISTRATION NUMBER: 35,570

REFERENCE/DOCKET NUMBER: 310068.402D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

TELEX: 3723836

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 223 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-143-579A-4

Query Match 99.4%; Score 498; DB 1; Length 223;

Best Local Similarity 98.9%; Pred. No. 7e-51;

Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPDQVKFPGGGQIVGGVLLPRRGPRLGVRATRKTSERSQPRG 60

Db 1 MSTNPKPQKTKRNTNRRPDQVKFPGGGQIVGGVLLPRRGPRLGVRATRKTSERSQPRG 60

QY 61 RROQIPKARRPGRRTWAQPGYWPPLYGNEG 91

Db 61 RROQIPKARRPGRRTWAQPGYWPPLYGNEG 91

RESULT 7

US-08-143-578A-4

Sequence 4, Application US/08143578A

Patent No. 5645983

GENERAL INFORMATION:

APPLICANT: Liao, Jaw-Ching

TITLE OF INVENTION: CORE ANTIGEN PROTEIN OF HEPATITIS C

TITLE OF INVENTION: VIRUS, AND DIAGNOSTIC METHOD AND KIT USING THE SAME

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: U.S.

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/143,578A

FILING DATE: 26-OCT-1993

CLASSIFICATION: 435

```

; ATTORNEY/AGENT INFORMATION:
; NAME: King, Joshua
; REGISTRATION NUMBER: 35,570
; REFERENCE/DOCKET NUMBER: 310075.402D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-143-578A-4

Query Match          99.4%; Score 498; DB 1; Length 223;
Best Local Similarity 98.9%; Pred. No. 7e-51;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQRTKRNTRRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
DB 1 MSTNPKPQRTKRNTRRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60

QY 61 RROPIKARRPEGRTWAQPGYPWPPLYGNEGM 91
DB 61 RROPIKARRPEGRTWAQPGYPWPPLYGNEGL 91

RESULT 8
US-08-454-928-8
; Sequence 8, Application US/08454928
; Patent No. 6153378
; GENERAL INFORMATION:
; APPLICANT: Liao, Jaw-Ching
; TITLE OF INVENTION: DIAGNOSIS OF, AND VACCINATION AGAINST, A
; TITLE OF INVENTION: POSITIVE STRANDED RNA VIRUS USING AN ISOLATED, UNPROCESSED
; TITLE OF INVENTION: POLYPEPTIDE ENCODED BY A SUBSTANTIALLY COMPLETE GENOME OF
; TITLE OF INVENTION: SUCH VIRUS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY
; STREET: 6300 Columbia Center
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,928
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: King, Joshua
; REGISTRATION NUMBER: 35,570
; REFERENCE/DOCKET NUMBER: 310075.405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836 SEEDANDBERRY
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-454-928-8

Query Match          99.4%; Score 498; DB 3; Length 223;
Best Local Similarity 98.9%; Pred. No. 1.2e-50;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

; ATTORNEY/AGENT INFORMATION:
; NAME: King, Joshua
; REGISTRATION NUMBER: 35,570
; REFERENCE/DOCKET NUMBER: 310075.402D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-143-578A-4

Query Match          99.4%; Score 498; DB 1; Length 223;
Best Local Similarity 98.9%; Pred. No. 7e-51;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQRTKRNTRRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
DB 1 MSTNPKPQRTKRNTRRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60

QY 61 RROPIKARRPEGRTWAQPGYPWPPLYGNEGM 91
DB 61 RROPIKARRPEGRTWAQPGYPWPPLYGNEGL 91

RESULT 9
US-08-462-195-2
; Sequence 2, Application US/08462195
; Patent No. 5789544
; GENERAL INFORMATION:
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: SAITO, IZUMU
; APPLICANT: MATSUURA, YOSHIHARU
; APPLICANT: HONDA, YOSHIKAZU
; TITLE OF INVENTION: METHOD FOR PRODUCING ECTOPROTEIN OF
; TITLE OF INVENTION: HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,195
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,303
; FILING DATE: 22-MAY-1995
; APPLICATION NUMBER: US 08/074,584
; FILING DATE: 11-JUN-1993
; APPLICATION NUMBER: JP 152487/1992
; FILING DATE: 11-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5789544man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 4169-003-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-462-195-2

Query Match          99.4%; Score 498; DB 1; Length 340;
Best Local Similarity 98.9%; Pred. No. 1.2e-50;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQRTKRNTRRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
DB 1 MSTNPKPQRTKRNTRRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
```

QY 61 RQPIKARRPGRRTWAQPGYWPPLYGNEG 91
Db 61 RQPIKARRPGRRTWAQPGYWPPLYGNEG 91

RESULT 10

US-08-636-883-2
; Sequence 2, Application US/08636883
; Patent No. 5830691
; GENERAL INFORMATION:
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: SAITO, IZUMU
; APPLICANT: MATSUURA, YOSHIMARU
; APPLICANT: HONDA, YOSHIKAZU
; APPLICANT: SEKI, MAKOTO
; TITLE OF INVENTION: METHOD FOR PRODUCING ECTOPROTEIN OF
; TITLE OF INVENTION: HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/636,883
; FILING DATE:
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,303
; FILING DATE:

; APPLICATION NUMBER: US 08/074,584
; FILING DATE: 11-JUN-1993
; APPLICATION NUMBER: JP 152487/1992
; FILING DATE: 11-JUN-1992
; ATTORNEY/AGENT INFORMATION:

; NAME: Oblon, No. 5830691man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 4169-003-0
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 340 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-08-636-883-2

Query Match 99.4%; Score 498; DB 2; Length 340;
Best Local Similarity 98.9%; Pred. No. 1.2e-50;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKFGGGQIVGGVLLPRGRLGVRATRKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKFGGGQIVGGVLLPRGRLGVRATRKTSERSQPRG 60

QY 61 RQPIKARRPGRRTWAQPGYWPPLYGNEG 91
Db 61 RQPIKARRPGRRTWAQPGYWPPLYGNEG 91

RESULT 11

US-09-127-829-2

; Sequence 2, Application US/09127829
; Patent No. 6063904
; GENERAL INFORMATION:
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: SAITO, IZUMU
; APPLICANT: MATSUURA, YOSHIMARU
; APPLICANT: HONDA, YOSHIKAZU
; APPLICANT: SEKI, MAKOTO
; TITLE OF INVENTION: METHOD FOR PRODUCING ECTOPROTEIN OF
; TITLE OF INVENTION: HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/127,829
; FILING DATE:
; CLASSIFICATION:

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,303
; FILING DATE:

; APPLICATION NUMBER: US 08/074,584
; FILING DATE: 11-JUN-1993
; APPLICATION NUMBER: JP 152487/1992
; FILING DATE: 11-JUN-1992

; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 6063904man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 4169-003-0
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 340 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-09-127-829-2

Query Match 99.4%; Score 498; DB 3; Length 340;
Best Local Similarity 98.9%; Pred. No. 1.2e-50;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKFGGGQIVGGVLLPRGRLGVRATRKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKFGGGQIVGGVLLPRGRLGVRATRKTSERSQPRG 60

QY 61 RQPIKARRPGRRTWAQPGYWPPLYGNEG 91
Db 61 RQPIKARRPGRRTWAQPGYWPPLYGNEG 91

RESULT 12

US-08-290-665A-172
; Sequence 172, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:

; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND

```

; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 172:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: HK3
US-08-290-665A-174

```

```

Query Match 99.2%; Score 497; DB 2; Length 191;
Best Local Similarity 98.9%; Pred. No. 7.7e-51;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPQRTKNTNRRPQDVKPGGGQIVGGVYLLPRGPRGLGVTRATKTSRSQPRG 60
Db 1 MSTNPKPQRTKNTNRRPQDVKPGGGQIVGGVYLLPRGPRGLGVTRATKTSRSQPRG 60

Qy 61 RROPIPKARQEGRTWAQPGYPWPPLYGNEGM 91
Db 61 RROPIPKARQEGRTWAQPGYPWPPLYGNEGM 91

RESULT 13
US-08-290-665A-174
; Sequence 174, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 174:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: HK4
US-08-290-665A-174

Query Match 99.2%; Score 497; DB 2; Length 191;
Best Local Similarity 98.9%; Pred. No. 7.7e-51;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPQRTKNTNRRPQDVKPGGGQIVGGVYLLPRGPRGLGVTRATKTSRSQPRG 60
Db 1 MSTNPKPQRTKNTNRRPQDVKPGGGQIVGGVYLLPRGPRGLGVTRATKTSRSQPRG 60

Qy 61 RROPIPKARQEGRTWAQPGYPWPPLYGNEGM 91
Db 61 RROPIPKARQEGRTWAQPGYPWPPLYGNEGM 91

RESULT 14
PCT-US95-10398-172
; Sequence 172, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10398
; FILING DATE: 15-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29 JUNE 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290/665

```

```
; FILING DATE: 15 AUGUST 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 172:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: HK3
PCT-US95-10398-172

Query Match          99.2%; Score 497; DB 5; Length 191;
Best Local Similarity 98.9%; Pred. No. 7.7e-51;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPDVKFPGGGQIVGGVYLLPRRGRLGVRATRKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPDVKFPGGGQIVGGVYLLPRRGRLGVRATRKTSERSQPRG 60

QY 61 RRQPIPKARRPEGRTWAQPGYWPPLYGNEGM 91
Db 61 RRQPIPKARQPEGRTWAQPGYWPPLYGNEGM 91

RESULT 15
PCT-US95-10398-174
; Sequence 174, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R. H. AND
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10398
; FILING DATE: 15-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29 JUNE 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290/665
; FILING DATE: 15 AUGUST 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
```

```
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 174:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; ORIGINAL SOURCE:
; ORGANISM: homosapiens
; INDIVIDUAL ISOLATE: HK4
PCT-US95-10398-174
```

```
Query Match          99.2%; Score 497; DB 5; Length 191;
Best Local Similarity 98.9%; Pred. No. 7.7e-51;
Matches 90; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPDVKFPGGGQIVGGVYLLPRRGRLGVRATRKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPDVKFPGGGQIVGGVYLLPRRGRLGVRATRKTSERSQPRG 60

QY 61 RRQPIPKARRPEGRTWAQPGYWPPLYGNEGM 91
Db 61 RRQPIPKARQPEGRTWAQPGYWPPLYGNEGM 91
```

Search completed: August 12, 2005, 14:24:47
Job time : 26.4889 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:04:14 ; Search time 18.1106 Seconds
(without alignments)
483.460 Million cell updates/sec

Title: US-09-758-308-1

Perfect score: 501

Sequence: 1 MSTNPKPQKTKRNTNRRPQ.....EGRTWAQPGYWPPLYGNEGM 91

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	501	100.0	115	2 S41343	genome polyprotein
2	501	100.0	3010	1 GNVVCJ	genome polyprotein
3	498	99.4	441	2 S12707	genome polyprotein
4	496	99.0	112	2 S41341	genome polyprotein
5	496	99.0	115	2 S41345	genome polyprotein
6	496	99.0	118	2 S41346	genome polyprotein
7	496	99.0	550	2 JH0711	genome polyprotein
8	496	99.0	3011	1 GNVVCH	genome polyprotein
9	494	98.6	115	2 S41342	genome polyprotein
10	492	98.2	115	2 S41351	genome polyprotein
11	492	98.2	640	2 JQ1584	genome polyprotein
12	492	98.2	3010	1 GNVVTC	genome polyprotein
13	490	97.8	115	2 S41344	genome polyprotein
14	489	97.6	782	2 S18031	hypothetical prote
15	489	97.6	787	2 PN0677	genome polyprotein
16	489	97.6	3010	1 S18030	structural protein
17	488	97.4	513	2 A44150	genome polyprotein
18	488	97.4	3011	1 GNVVC3	genome polyprotein
19	487	97.2	115	2 S41350	genome polyprotein
20	487	97.2	115	2 S41347	genome polyprotein
21	487	97.2	189	2 S32740	polyprotein - hepa
22	487	97.2	369	2 S21471	genome polyprotein
23	485	96.8	513	2 PC1284	genome polyprotein
24	485	96.8	520	2 JQ1925	polyprotein - hepa
25	485	96.8	523	2 JQ1926	polyprotein - hepa
26	485	96.8	782	2 S19876	genome polyprotein
27	483	96.4	114	2 S41362	genome polyprotein
28	483	96.4	3011	1 S40770	genome polyprotein
29	479	95.6	114	2 S41366	genome polyprotein

ALIGNMENTS

RESULT 1

S41343 genome polyprotein - hepatitis C virus (genotype 1, N11) (fragment)

N:Contains: core protein
C:Species: hepatitis C virus

A:Variety: genotype 1, N11

C:Date: 19-May-1994 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
Accession: S41343

R:van Doorn, L.J.; Kletter, G.E.M.; Brouwer, J.T.

submitted to the EMBL Data Library, January 1994

A:Description: Analysis of hepatitis C virus genotypes 1 to 5 by LiPA.

A:Reference number: S41341

A:Accession: S41343

A:Molecule type: genomic RNA

A:Residues: 1-115 <VAR>

A:Cross-references: UNIPROT:Q68877; EMBL:229446; NID:G443854; PIDN:CAA82584.1; PID:G443.

A:Experimental source: genotype 1, N11

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: capsid protein; core protein; polyprotein

F1-115/Product: core protein #status predicted <MAT>

Query Match 100.0%; Score 501; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. No. 8.8e-42;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDKVFGGQIVGGVLLPRRGPRGLGVRATKTSERSQPRG 60

Db 1 MSTNPKPQKTKRNTNRRPQDKVFGGQIVGGVLLPRRGPRGLGVRATKTSERSQPRG 60

QY 61 RROPIPKARRPGRGRTWAQPGYWPPLYGNEGM 91

Db 61 RROPIPKARRPGRGRTWAQPGYWPPLYGNEGM 91

RESULT 2

GNVVCJ

genome polyprotein - hepatitis C virus (strain J)

N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

C:Species: hepatitis C virus

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004

A:Accession: A39253; PS0086

R:Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimo

Proc. Natl. Acad. Sci. U.S.A. 87, 5524-5528, 1990

A:Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients

A:Reference number: A39253; MUID:91088550; PMID:2175903

A:Accession: A39253

A:Molecule type: genomic RNA

A:Residues: 1-3010 <VAR>

A:Cross-references: UNIPROT:P26662; GB:D90208; NID:G221610; PIDN:BAAL4233.1; PID:G22161

R:Kato, N.; Ohkoshi, S.; Shimotohno, K.

Proc. Jpn. Acad. 65B, 219-223, 1989

Db 61 RROPIPKARRPEGRRTWAQPGYFWPLYGNEG 90

RESULT 14

S18031
genome polyprotein - hepatitis C virus (isolate JK2) (fragment)
N:Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein
C:Species: hepatitis C virus
A:Variety: isolate JK2
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: S18031
R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A:Description: Sequence analysis of putative structural regions of Hepatitis C Virus iso
A:Reference number: S18029
A:Accession: S18031
A:Molecule type: genomic RNA
A:Residues: 1-782 <HON>
A:Cross-references: UNIPROT:Q68950; EMBL:X61593
A:Experimental source: isolate JK2
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural
F:1-191/Product: core protein #status predicted <MAT1>
F:192-383/Product: envelope protein 1 #status predicted <MAT2>
F:384-733/Product: NS1/E2 protein #status predicted <MAT3>
F:734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Query Match 97.6%; Score 489; DB 2; Length 782;
Best Local Similarity 96.7%; Pred. No. 8e-40;
Matches 88; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVILLPRGPRGLGVRATRKTSERSQPRG 60
Db |||||||
QY 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVILLPRGPRGLGVRATRKTSERSQPRG 60
Db |||||||

QY 61 RROPIPKARRPEGRRTWAQPGYFWPLYGNEG 91
Db |||||||

RESULT 15

PN0677
hypothetical protein 787 - hepatitis C virus (fragment)
C:Species: hepatitis C virus
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
C:Accession: PN0677
R:Cho, S.H.; Yoon, J.I.; Chang, J.E.; Ahn, B.M.; Lee, C.H.; Lee, Y.I.
Biochem. Biophys. Res. Commun. 196, 780-788, 1993
A:Title: Genomic typing of hepatitis C viruses from Korean patients: Implications of gen
A:Reference number: PN0677; MUID:94059104; PMID:8240354
A:Accession: PN0677
A:Molecule type: mRNA
A:Residues: 1-787 <CHO>
A:Cross-references: UNIPROT:Q08244; GB:I20498; NID:gl381031; PIDN:AAB02608.1; PID:gl3810
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: glycoprotein; nonstructural protein
F:196,209,233,250,305,325,421,427,452,536,544,560,580,627,649/Binding site: carbohydrate

Query Match 97.6%; Score 489; DB 2; Length 787;
Best Local Similarity 96.7%; Pred. No. 8e-40;
Matches 88; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVILLPRGPRGLGVRATRKTSERSQPRG 60
Db |||||||
QY 61 RROPIPKARRPEGRRTWAQPGYFWPLYGNEG 91
Db |||||||

Search completed: August 12, 2005, 14:22:46
Job time : 19.1106 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:03:18 ; Search time 84.0688 Seconds
(without alignments)
554.298 Million cell updates/sec

Title: US-09-758-308-1

Perfect score: 501

Sequence: 1 MSTNPKPQKTKRNTNRRPQ.....EGRTWAQGYWPPLYGNEGM 91

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	501	100.0	100	2	Q8QP58 hepatitis c
2	501	100.0	100	2	Q8QP60 hepatitis c
3	501	100.0	100	2	Q8QP65 hepatitis c
4	501	100.0	115	2	Q68877 hepatitis c
5	501	100.0	135	2	Q70GDD hepatitis c
6	501	100.0	191	2	Q8JWL3 hepatitis c
7	501	100.0	191	2	Q8JWN0 hepatitis c
8	501	100.0	191	2	Q8JWN3 hepatitis c
9	501	100.0	191	2	Q8V7J3 hepatitis c
10	501	100.0	191	2	Q8V7P2 hepatitis c
11	501	100.0	191	2	Q8V7P3 hepatitis c
12	501	100.0	191	2	Q8V7P4 hepatitis c
13	501	100.0	191	2	Q8V7P5 hepatitis c
14	501	100.0	191	2	Q8V7P6 hepatitis c
15	501	100.0	191	2	Q8V7P7 hepatitis c
16	501	100.0	191	2	Q8V7P8 hepatitis c
17	501	100.0	191	2	Q8V7P9 hepatitis c
18	501	100.0	191	2	Q8V7Q0 hepatitis c
19	501	100.0	191	2	Q8V7Q1 hepatitis c
20	501	100.0	191	2	Q8V7Q2 hepatitis c
21	501	100.0	191	2	Q8V7Q4 hepatitis c
22	501	100.0	191	2	Q8V7Q5 hepatitis c
23	501	100.0	191	2	Q8V7Q6 hepatitis c
24	501	100.0	191	2	Q8V7Q8 hepatitis c
25	501	100.0	191	2	Q8V7Q9 hepatitis c
26	501	100.0	191	2	Q8V7R0 hepatitis c
27	501	100.0	191	2	Q8V7R1 hepatitis c
28	501	100.0	191	2	Q8V7Z3 hepatitis c
29	501	100.0	191	2	Q8V7Z4 hepatitis c
30	501	100.0	191	2	Q8V7Z6 hepatitis c
31	501	100.0	191	2	Q8V7Z7 hepatitis c

32 501 100.0 191 2 Q8V7Z8 hepatitis c
33 501 100.0 191 2 Q8V7Z9 hepatitis c
34 501 100.0 191 2 Q8V800 hepatitis c
35 501 100.0 191 2 Q8V801 hepatitis c
36 501 100.0 191 2 Q8V803 hepatitis c
37 501 100.0 191 2 Q8V804 hepatitis c
38 501 100.0 191 2 Q8V805 hepatitis c
39 501 100.0 191 2 Q8V806 hepatitis c
40 501 100.0 191 2 Q8V807 hepatitis c
41 501 100.0 191 2 Q8V808 hepatitis c
42 501 100.0 191 2 Q8V809 hepatitis c
43 501 100.0 191 2 Q8V810 hepatitis c
44 501 100.0 191 2 Q8V811 hepatitis c
45 501 100.0 191 2 Q8V812 hepatitis c

ALIGNMENTS

RESULT 1

Q8QP58 PRELIMINARY; PRT; 100 AA.
AC Q8QP58;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=783;
RX MEDLINE=21904745; PubMed=11907242;
RX DOI=10.1128/JVI.76.8.4034-4043.2002;
RA Kalinina O., Norder H., Mukomolov S., Magnus L.O.;
RT "A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg."
RL J. Virol. 76:4034-4043(2002).
DR EMBL; AY070207; AAL58621.1; -.
DR FIR; PQ0804; PQ0804.
DR HSP; Q8JYSL; ICWX.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV_capsid; 1.
FT NON_TER 100 100
SQ SEQUENCE 100 AA; 11357 MW; E211A2C13010C48D CRC64;

Query Match 100.0%; Score 501; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.1e-43;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGRLGVRAATKTSERSQPRG 60
|||||
Db 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGRLGVRAATKTSERSQPRG 60
|||||

QY 61 RROPIPKARRPEGRTWAQGYWPPLYGNEGM 91
|||||

Db 61 RROPIPKARRPEGRTWAQGYWPPLYGNEGM 91
|||||

RESULT 2

Q8QP60 PRELIMINARY; PRT; 100 AA.
ID Q8QP60;
AC Q8QP60;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC

```

OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=788;
RX MEDLINE=21904745; PubMed=11907242;
RX DOI=10.1128/JVI.76.8.4034-4043.2002;
RT "A natural intergenotypic recombinant of hepatitis C virus identified
RT in St. Petersburg.";
RL J. Virol. 76:4034-4043(2002).
DR EMBL; AY070205; AAL58619.1; -.
DR PIR; PQ0804; PQ0804.
DR HSSP; Q8JYS1; 1CWK.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV capsid; 1.
FT NON_TER 100 100
SQ SEQUENCE 100 AA; 11357 MW; E211A2C13010C48D CRC64;

Query Match 100.0%; Score 501; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.1e-43;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60

Qy 61 RROPIPKARREGRTWAQPGYPWPPLYGNEG 91
Db 61 RROPIPKARREGRTWAQPGYPWPPLYGNEG 91

RESULT 3
Qy Q8QP65 PRELIMINARY; PRT; 100 AA.
AC Q8QP65;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=732;
RX MEDLINE=21904745; PubMed=11907242;
RX DOI=10.1128/JVI.76.8.4034-4043.2002;
RT "A natural intergenotypic recombinant of hepatitis C virus identified
RT in St. Petersburg.";
RL J. Virol. 76:4034-4043(2002).
DR EMBL; AY070200; AAL58614.1; -.
DR PIR; PQ0804; PQ0804.
DR HSSP; Q8JYS1; 1CWK.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV capsid; 1.
FT NON_TER 100 100
SQ SEQUENCE 100 AA; 11357 MW; E211A2C13010C48D CRC64;

Query Match 100.0%; Score 501; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.1e-43;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60

Qy 61 RROPIPKARREGRTWAQPGYPWPPLYGNEG 91
Db 61 RROPIPKARREGRTWAQPGYPWPPLYGNEG 91

RESULT 5
Qy Q70GDO PRELIMINARY; PRT; 135 AA.
AC Q70GDO;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus type 1.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=41856;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14718617; DOI=10.1099/vir.0.19472-0;
RA Colina R., Casane D., Vasquez S., Garcia L., Chunga A., Romero H.,

```

```

Qy 61 RROPIPKARREGRTWAQPGYPWPPLYGNEG 91
Db 61 RROPIPKARREGRTWAQPGYPWPPLYGNEG 91

RESULT 4
Qy Q68877 PRELIMINARY; PRT; 115 AA.
AC Q68877;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE (genotype 1, B75) core region RNA (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97201609; PubMed=9049395;
RX van Doorn L.J., Kleter B.G.E.M., Stuyver L., Maertens G.,
RA Brouwer J.T., Schalm S.W., Heijtkink R.A., Quint W.G.V.;
RT "Sequence analysis of Hepatitis C virus genotypes 1 to 5 reveals
RT multiple novel subtypes in the Benelux countries.";
RL J. Gen. Virol. 76:1871-1876(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96048319; PubMed=8551256;
RX Kleter G.E.M., van Doorn L.J., Stuyver L., Maertens G., Brouwer J.T.,
RA Schalm S.W., Heijtkink R.A., Quint W.G.V.;
RT "Rapid genotyping of hepatitis C virus RNA-isolates obtained from
RT patients residing in western Europe.";
RL J. Med. Virol. 47:35-42(1995).
DR EMBL; Z29446; CAA82584.1; -.
DR PIR; PQ0804; PQ0804.
DR PIR; S41343; S41343.
DR HSSP; Q8JYS1; 1CWK.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV capsid; 1.
FT NON_TER 115 115
SQ SEQUENCE 115 AA; 13120 MW; 6304D6F9529F138A CRC64;

Query Match 100.0%; Score 501; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. No. 3.6e-43;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60

Qy 61 RROPIPKARREGRTWAQPGYPWPPLYGNEG 91
Db 61 RROPIPKARREGRTWAQPGYPWPPLYGNEG 91

RESULT 5
Qy Q70GDO PRELIMINARY; PRT; 135 AA.
AC Q70GDO;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus type 1.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=41856;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14718617; DOI=10.1099/vir.0.19472-0;
RA Colina R., Casane D., Vasquez S., Garcia L., Chunga A., Romero H.,

```



```
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
KW Polyprotein.
FT NON_TER 1
FT CHAIN <1>191 core protein.
FT NON_TER 191
SQ SEQUENCE 191 AA; 20796 MW; 93DD39B92A7E8031 CRC64;

Query Match 100.0%; Score 501; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 6.2e-43;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKFFGGQIVGGVYLLPRRGLGVRAVTRTKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKFFGGQIVGGVYLLPRRGLGVRAVTRTKTSERSQPRG 60

QY 61 RROPIPKARRPEGRRTWAQPGYPWPLYGNEGM 91
Db 61 RROPIPKARRPEGRRTWAQPGYPWPLYGNEGM 91

RESULT 9
Q8V7J3 PRELIMINARY; PRT; 191 AA.
AC Q8V7J3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Core protein (Fragment).
GN Name=core;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Kato N.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB063486; BAB83519.1; -.
DR PIR; PQ0804; PQ0804.
DR HSP; Q8JYS1; ICWX.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON_TER 191
SQ SEQUENCE 191 AA; 20808 MW; 93DD39B93B8E8031 CRC64;

Query Match 100.0%; Score 501; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 6.2e-43;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKFFGGQIVGGVYLLPRRGLGVRAVTRTKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKFFGGQIVGGVYLLPRRGLGVRAVTRTKTSERSQPRG 60

QY 61 RROPIPKARRPEGRRTWAQPGYPWPLYGNEGM 91
Db 61 RROPIPKARRPEGRRTWAQPGYPWPLYGNEGM 91

RESULT 10
Q8V7P2 PRELIMINARY; PRT; 191 AA.
AC Q8V7P2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Core protein (Fragment).
GN Name=core;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Kato N.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062232; BAB83426.1; -.
DR PIR; PQ0804; PQ0804.
DR HSP; Q8JYS1; ICWX.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON_TER 191
SQ SEQUENCE 191 AA; 20808 MW; 93DD39B93B8E8031 CRC64;

Query Match 100.0%; Score 501; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 6.2e-43;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKFFGGQIVGGVYLLPRRGLGVRAVTRTKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKFFGGQIVGGVYLLPRRGLGVRAVTRTKTSERSQPRG 60

QY 61 RROPIPKARRPEGRRTWAQPGYPWPLYGNEGM 91
Db 61 RROPIPKARRPEGRRTWAQPGYPWPLYGNEGM 91

RESULT 11
Q8V7P3 PRELIMINARY; PRT; 191 AA.
AC Q8V7P3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Core protein (Fragment).
GN Name=core;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Kato N.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062231; BAB83425.1; -.
DR PIR; PQ0804; PQ0804.
DR HSP; Q8JYS1; ICWX.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON_TER 191
SQ SEQUENCE 191 AA; 20774 MW; 9D33D9B93B848A31 CRC64;

Query Match 100.0%; Score 501; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 6.2e-43;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKFFGGQIVGGVYLLPRRGLGVRAVTRTKTSERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKFFGGQIVGGVYLLPRRGLGVRAVTRTKTSERSQPRG 60

QY 61 RROPIPKARRPEGRRTWAQPGYPWPLYGNEGM 91
Db 61 RROPIPKARRPEGRRTWAQPGYPWPLYGNEGM 91
```



```
RESULT 12
Q8V7P4
ID Q8V7P4 PRELIMINARY; PRT; 191 AA.
AC Q8V7P4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Core protein (Fragment).
GN Name=core;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Kato N.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062230; BAB83424.1; -.
DR PIR; PQ0804; PQ0804.
DR HSSP; Q8JYS1; ICWX.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
FT NON_TER 191
SQ SEQUENCE 191 AA; 20808 MW; 93DD39B93B8E8031 CRC64;

Query Match 100.0%; Score 501; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 6.2e-43;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60
DB 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60

QY 61 RRQIPKARRPEGRTWAQPGYWPPLYGNEGM 91
DB 61 RRQIPKARRPEGRTWAQPGYWPPLYGNEGM 91

RESULT 13
Q8V7P5
ID Q8V7P5 PRELIMINARY; PRT; 191 AA.
AC Q8V7P5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Core protein (Fragment).
GN Name=core;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Kato N.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062229; BAB83423.1; -.
DR PIR; PQ0804; PQ0804.
DR HSSP; Q8JYS1; ICWX.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
FT NON_TER 191
SQ SEQUENCE 191 AA; 20808 MW; 93DD39B93B8E8031 CRC64;

Query Match 100.0%; Score 501; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 6.2e-43;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60
DB 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60

QY 61 RRQIPKARRPEGRTWAQPGYWPPLYGNEGM 91
DB 61 RRQIPKARRPEGRTWAQPGYWPPLYGNEGM 91

RESULT 14
Q8V7P6
ID Q8V7P6 PRELIMINARY; PRT; 191 AA.
AC Q8V7P6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Core protein (Fragment).
GN Name=core;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Kato N.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062228; BAB83422.1; -.
DR PIR; PQ0804; PQ0804.
DR HSSP; Q8JYS1; ICWX.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
FT NON_TER 191
SQ SEQUENCE 191 AA; 20808 MW; 93DD39B93B8E8031 CRC64;

Query Match 100.0%; Score 501; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 6.2e-43;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60
DB 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRATRKTSERSQPRG 60

QY 61 RRQIPKARRPEGRTWAQPGYWPPLYGNEGM 91
DB 61 RRQIPKARRPEGRTWAQPGYWPPLYGNEGM 91

RESULT 15
Q8V7P7
ID Q8V7P7 PRELIMINARY; PRT; 191 AA.
AC Q8V7P7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Core protein (Fragment).
GN Name=core;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Kato N.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062227; BAB83421.1; -.
DR PIR; PQ0804; PQ0804.
DR HSSP; Q8JYS1; ICWX.
DR GO; GO:0019028; C:Viral capsid; IEA.
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:00:18 ; Search time 82.8821 Seconds
(without alignments)
368.645 Million cell updates/sec

Title: US-09-758-308-3

Perfect score: 394

Sequence: 1 SVVIVGRILSRPAVDP.....TATKQAEAAAPVVEKRWAL 79

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	394	100.0	79	AAB31697	Aab31697 Antigenic
2	394	100.0	194	AAR52737	Aar52737 HCV antig
3	394	100.0	195	AAR25863	Aar25863 HCV poly
4	394	100.0	195	AAW41741	Aaw41741 Hepatitis
5	394	100.0	293	AAR98350	Aar98350 HCV NS4 d
6	394	100.0	477	AAR29866	Aar29866 HCV NS2-N
7	394	100.0	477	AAR29867	Aar29867 HCV NS2-N
8	394	100.0	477	AAR29865	Aar29865 HCV NS2-N
9	394	100.0	768	AAR29868	Aar29868 HCV NS2-N
10	394	100.0	768	AAR29850	Aar29850 HCV NS2-N
11	394	100.0	768	AAR29869	Aar29869 HCV NS2-N
12	394	100.0	1188	AAR29660	Aar29660 HCV NS2-N
13	394	100.0	2510	AAR29527	Aar29527 HCV antig
14	394	100.0	3010	AAR68864	Aar68864 Hepatitis
15	394	100.0	3010	AAR82694	Aar82694 Partial H
16	394	100.0	3010	ADP88597	Adp88597 Hepatitis
17	392	99.5	373	AAR41436	Aar41436 PT-NANBH
18	392	99.5	1250	AAR12599	Aar12599 Portion o
19	392	99.5	2354	AAR41435	Aar41435 PT-NANBH
20	391	99.2	3010	AAR68622	Aar68622 HCV prote
21	390	99.0	1411	AAR29533	Aar29533 HCV NS4-N
22	390	99.0	3010	AAR34580	Aar34580 Human hep
23	389	98.7	86	AAW09051	Aaw09051 Hepatitis
24	389	98.7	200	AAR29846	Aar29846 HCV NS2-N
25	389	98.7	1736	AAB36932	Aab36932 Hepatitis

ALIGNMENTS

RESULT 1

AAB31697

ID AAB31697 standard; peptide; 79 AA.

XX AC AAB31697;

XX DT 30-APR-2001 (first entry)

XX DE Antigenic epitope of the Hepatitis C virus (HCV) NS4 protein.

XX KW Antigen; HCV; polyprotein; core protein; NS3 protein; NS4 protein;

XX KW NS4a protein; HCV infection.

XX OS Hepatitis C virus.

XX PN WO200104149-A1.

XX PD 18-JAN-2001.

XX PF 07-JUL-2000; 2000WO-US018704.

XX PR 09-JUL-1999; 99WO-US015578.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Fields HA, Khudyakov YE;

XX DR WPI; 2001-138316/14.

XX PT New (mosaic) polypeptides, useful as reagents in assays for the diagnosis or monitoring of HCV infections and as components of anti-HCV vaccines, comprises antigenic groups of hepatitis C virus (HCV).

XX PS Claim 4; Page 39; 52pp; English.

XX CC The present sequence represents an antigenic epitope from a hepatitis C virus (HCV) NS4 protein. The peptide comprises amino acids 1789-1867 of the HCV polyprotein. The specification describes antigenic epitopes from HCV core protein, NS3 protein, NS4 protein and NS4a protein. The antigenic peptides are useful as diagnostic reagents for detecting HCV in a biological sample. They are also useful for monitoring HCV infection in a patient sample in addition to diagnosis. Pharmaceutical compositions comprising the peptides are useful for preventing, minimizing or reducing HCV infection in patients who have been exposed to HCV or to individuals, such as health care workers or blood product recipients, who are more likely to become exposed to HCV infection. The antigenic peptides are also useful for generating antibodies which can be used to detect HCV proteins in a sample or for laboratory research purposes

Aaw01680 HCV NS2-N
Aay70064 Recombina
Aay70065 Recombina
Aay70066 Recombina
Aar20111 Non-A, no
Aar20091 Non-A, no
Aay06423 Non-A, no
Add67962 EMCV inte
Abu09574 HCV Met-N
Abu09575 HCV Met-N
Aar38450 Hepatitis
Aae20477 HCV-S1 fu
Adi95303 OSPP-rela
Aar30616 Polypepti
Aar53417 Blood tra
Aar49652 HCV pepti
Aar29907 HCV anti
Aar58591 Hepatitis
Adl17782 Hepatitis
Aar34468 Encoded b

26 389 98.7 2201 2 AAW01680
27 389 98.7 2307 3 AAY70064
28 389 98.7 2307 3 AAY70065
29 389 98.7 2307 3 AAY70066
30 389 98.7 3010 2 AAR20111
31 389 98.7 3010 2 AAR20091
32 389 98.7 3010 2 AAY06423
33 389 98.7 3090 7 ADD67962
34 387 98.2 1985 6 ABU09574
35 387 98.2 1985 6 ABU09575
36 387 98.2 1985 8 ADR38450
37 387 98.2 3010 5 AAE20477
38 386 98.0 2280 8 ADI95303
39 386 98.0 3010 2 AAR30616
40 386 98.0 3010 2 AAR53417
41 385 97.7 87 2 AAR49652
42 385 97.7 320 2 AAR29907
43 385 97.7 697 2 AAR58591
44 385 97.7 697 8 ADL17782
45 385 97.7 3011 2 AAR34468

```

XX SQ Sequence 79 AA;
Query Match 100.0%; Score 394; DB 4; Length 79;
Best Local Similarity 100.0%; Pred. No. 4.4e-42;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVVIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLAEQFKQKALGLLOT 60
Db 1 SVVIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLAEQFKQKALGLLOT 60

Qy 61 ATKQAEAAAPVVESKWRAL 79
Db 61 ATKQAEAAAPVVESKWRAL 79

RESULT 2
AAR52737
ID AAR52737 standard; protein; 194 AA.
AC AAR52737;
XX 31-JAN-1995 (first entry)
XX HCV antigen.
XX Hepatitis C virus; HCV; antigen; diagnosis; reagent; agglutination.
XX Synthetic.
XX JP06102273-A.
XX 15-APR-1994.
XX 18-SEP-1992; 92JP-00250027.
XX 18-SEP-1992; 92JP-00250027.
XX (TOKU ) TOKUYAMA SODA KK.
XX WPI; 1994-161280/20.
XX N-PSDB; AAQ62690.
XX Immunological agglutination reagent for the diagnosis of hepatitis C -
XX comprising hepatitis C virus antigen polypeptide.
XX Claim 2-3; Page 17-18; 18pp; Japanese.
XX An new immunological agglutination reaction reagent for the diagnosis of
XX hepatitis C uses a HCV antigen polypeptide subjected to heat-treatment.
XX The HCV antigen active polypeptide contains one of the polypeptides given
XX in AAR52735-38
XX Sequence 194 AA;
Query Match 100.0%; Score 394; DB 2; Length 194;
Best Local Similarity 100.0%; Pred. No. 1.5e-41;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVVIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLAEQFKQKALGLLOT 60
Db 75 SVVIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLAEQFKQKALGLLOT 134

Qy 61 ATKQAEAAAPVVESKWRAL 79
Db 135 ATKQAEAAAPVVESKWRAL 153

RESULT 3
AAR25863
ID AAR25863 standard; protein; 195 AA.
XX AAR25863;
AC AAR25863;

```

```

XX 21-JAN-1993 (first entry)
DT HCV polypeptide 10.
XX Recombinant vector; E. coli; diagnostic; reagent; type C hepatitis.
XX Hepatitis C virus.
XX JP04179482-A.
XX 26-JUN-1992.
XX 11-NOV-1990; 90JP-00304417.
XX 11-NOV-1990; 90JP-00304417.
XX (TOKU ) TOKUYAMA SODA KK.
XX WPI; 1992-263663/32.
XX N-PSDB; AAQ26990.
XX Hepatitis C virus antigen expressed as recombinant in E.coli - useful for
XX diagnosis of hepatitis C virus infection.
XX Claim 1; Page 3-4; 66pp; Japanese.
XX The sequences given in AAR25854-74 are hepatitis C virus proteins. The
XX genes encoding these proteins can each be used to prepare recombinant
XX vectors by ligating the gene of interest in to a vector to be expressed
XX in E. coli. These polypeptides are useful as diagnostic reagents for type
XX C hepatitis and they may be produced efficiently by recombinant methods
XX Sequence 195 AA;
Query Match 100.0%; Score 394; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.5e-41;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVVIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLAEQFKQKALGLLOT 60
Db 76 SVVIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLAEQFKQKALGLLOT 135

Qy 61 ATKQAEAAAPVVESKWRAL 79
Db 136 ATKQAEAAAPVVESKWRAL 154

RESULT 4
AAR41741
ID AAR41741 standard; protein; 195 AA.
XX AAR41741;
XX 22-MAY-1998 (first entry)
XX Hepatitis C virus antigen.
XX non-A non-B hepatitis virus; NANBH; hepatitis C virus; HCV; antigen;
XX diagnosis; detection.
XX Hepatitis virus.
XX JP05176774-A.
XX 20-JUL-1993.
XX 18-DEC-1991; 91JP-00354708.
XX 18-DEC-1990; 90JP-00412020.
XX (SHIM/) SHIMOTONO K.
XX (GREC ) GREEN CROSS CORP.
PA

```

XX WPI; 1993-260858/33.
DR Protein contg. non-A non-B hepatitis antigen fragment - prepd. by
PT culturing transformants transformed by vector contg. base sequence coding
PT specified aminoacid sequences, used for detecting hepatitis.
XX
PS Claim 1; Fig 7; 53pp; Japanese.
XX
CC The present sequence is a non-A non-B hepatitis virus (NANBH) or
CC hepatitis C virus (HCV) antigen, useful for diagnosis or detection
XX
SQ Sequence 195 AA;

Query Match 100.0%; Score 394; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.5e-41;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVVIVGRILSGRPAVIPDREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLQT 60
Db 76 SVVIVGRILSGRPAVIPDREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLQT 135

QY 61 ATKQAEAAAPVVESKWRAL 79
Db 136 ATKQAEAAAPVVESKWRAL 154

RESULT 5
AAR98350
ID AAR98350 standard; peptide; 293 AA.
AC AAR98350;
XX
DT 21-NOV-1996 (first entry)
XX
DE HCV NS4 derived antigen.
XX
KW Antigen; NS4; hepatitis C virus; HCV; antibody; bird; egg yolk;
KW detection; NS3; blood; diagnosis.
XX
OS Hepatitis C virus.
XX
XX JP08127598-A.
PN
XX
PD 21-MAY-1996.
XX
PF 28-OCT-1994; 94JP-00264808.
XX
PR 28-OCT-1994; 94JP-00264808.
XX
XX (SHIE) SHINETSU CHEM IND CO LTD.
PA (MITU) MITSUBISHI CHEM CORP.
XX
XX WPI; 1996-295556/30.
XX
XX Anti-HCV NS3 and NS4 protein antibodies - produced by immunising a bird
PT and isolating the antibodies from egg yolk.
PT
XX
PS Claim 2; Page 7-8; 9pp; English.
XX
CC The sequences given in AAR98349-50 represent antigenic peptides which
CC were derived from the NS3 and NS4 proteins of hepatitis C virus (HCV),
CC respectively. These peptides were used in the preparation of novel
CC antibodies. The antibodies were prepared by immunising a bird with one of
CC the peptides and isolating the resulting antibody from the egg yolk of
CC eggs laid by the immunised bird. The antibodies may be used in the
CC detection of HCV antigens in the blood and are useful in the development
CC of diagnostic agents or drugs for hepatitis C. The antibody is highly
CC specific and as the peptide is injected without a carrier protein, only
CC anti-NS3 or anti-NS4 antibodies are produced
XX
SQ Sequence 293 AA;

Query Match 100.0%; Score 394; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 2.5e-41;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVVIVGRILSGRPAVIPDREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLQT 60
Db 116 SVVIVGRILSGRPAVIPDREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLQT 175

QY 61 ATKQAEAAAPVVESKWRAL 79
Db 176 ATKQAEAAAPVVESKWRAL 194

RESULT 6
AAR29866
ID AAR29866 standard; protein; 477 AA.
XX
AC AAR29866;
XX
DT 25-MAR-2003 (revised)
DT 26-APR-1993 (first entry)
XX
DE HCV NS2-NS4 peptide N16N15B-1.
XX
KW Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum; HC;
KW transcriptase; cDNA; primer; allele.
XX
OS Hepatitis C virus.
XX
XX EP518313-A2.
PN
XX
PD 16-DEC-1992.
XX
PF 11-JUN-1992; 92EP-00109812.
XX
PR 11-JUN-1991; 91JP-00139268.
PR 12-JUL-1991; 91JP-00172794.
PR 07-OCT-1991; 91JP-00287008.
PR 16-DEC-1991; 91JP-00332329.
PR 20-APR-1992; 92JP-00099957.
XX
XX (MITU) MITSUBISHI KASEI CORP.
XX
XX Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayaashi N;
PI
XX WPI; 1992-417213/51.
DR N-PSDB; AAQ32497.
XX
PT New hepatitis C virus gene and its encoded protein - used for diagnosing
PT and vaccinating against hepatitis C virus infections.
XX
PS Disclosure; Page 172-75; 305pp; English.
XX
CC The sequences given in AAR29852-70 are encoded by various clones which
CC were used in the isolation of the NS2-NS4 regions of the Hepatitis C
CC Virus (HCV) gene of the invention (see also AAR29660, AAR29559-60 and
CC AAR29843-51). These RNA sequences were isolated from the serum of a
CC patient suffering from hepatitis C (HC). The isolated RNA sequences were
CC converted into cDNA using transcriptase in the presence of one of the
CC primer sequences given in AAQ32578-79. The sequences were then amplified
CC using primer pairs. The cDNA sequences isolated represent different
CC alleles of the same region of the HCV gene. Sequence comparisons of these
CC clones showed that it is possible for a patient to carry more than one
CC HCV strain at one time. See also AAQ32436. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 477 AA;

Query Match 100.0%; Score 394; DB 2; Length 477;
Best Local Similarity 100.0%; Pred. No. 4.8e-41;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVVIVGRILSGRPAVIPDREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLQT 60

ID AAR29868 standard; protein; 768 AA.
XX AC AAR29868;
XX XX
XX 25-MAR-2003 (revised)
DT 26-APR-1993 (first entry)
XX DT
XX HCV NS2-NS4 peptide N23N15A-1.
DE
XX Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum; HC;
KW transcriptase; cDNA; primer; allele.
XX
XX Hepatitis C virus.
XX
XX Key Location/Qualifiers
FH Misc-difference 19 /label= Pro, Leu
FT
FT Misc-difference 24 /label= His, Arg
FT
FT Misc-difference 60 /label= Thr, Ala
FT
FT Misc-difference 71 /label= Ser, Thr
FT
FT Misc-difference 88 /label= Lys, Arg
FT
FT Misc-difference 140 /label= Ile, Val
FT
FT Misc-difference 241 /label= Thr, Ile
FT
FT Misc-difference 274 /label= Val, Ile
FT
FT Misc-difference 381 /label= Pro, Ser
FT
FT Misc-difference 475 /label= Tyr, His
FT
FT Misc-difference 483 /label= Gln, Lys
FT
FT Misc-difference 485 /label= Thr, Ala
FT
FT Misc-difference 493 /label= Tyr, Phe
FT
FT Misc-difference 496 /label= Thr, Ala
FT
XX EP518313-A2.
XX PN
XX 16-DEC-1992.
XX PD
XX 11-JUN-1992; 92EP-00109812.
XX PF
XX 11-JUN-1991; 91JP-00139268.
XX PR
XX 12-JUL-1991; 91JP-00172794.
XX PR
XX 07-OCT-1991; 91JP-00287008.
XX PR
XX 16-DEC-1991; 91JP-00332329.
XX PR
XX 20-APR-1992; 92JP-00099957.
XX PR
XX (MITU) MITSUBISHI KASEI CORP.
XX PA
XX Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayaishi N;
XX PI
XX WPI; 1992-417213/51.
XX DR
XX N-PSDB; AAQ32499.
XX DR
XX New hepatitis C virus gene and its encoded protein - used for diagnosing
PT and vaccinating against hepatitis C virus infections.
XX PT
XX Disclosure; Page 178-82; 305pp; English.
XX PS
XX The sequences given in AAR29852-70 are encoded by various clones which
CC were used in the isolation of the NS2-NS4 regions of the Hepatitis C
CC Virus (HCV) gene of the invention (see also AAR29660, AAR29559-60 and
CC AAR29843-51). These RNA sequences were isolated from the serum of a
CC patient suffering from hepatitis C (HC). The isolated RNA sequences were
CC converted into cDNA using transcriptase in the presence of one of the
CC primer sequences given in AAQ32578-79. The sequences were then amplified
CC using primer pairs. The cDNA sequences isolated represent different
CC alleles of the same region of the HCV gene. Sequence comparisons of these
CC clones showed that it is possible for a patient to carry more than one
CC HCV strain at one time. See also AAQ32436. (Updated on 25-MAR-2003 to
XX correct PN field.)
XX CC
XX Sequence 768 AA;
XX SQ
Query Match 100.0%; Score 394; DB 2; Length 768;
Best Local Similarity 100.0%; Pred. No. 9.1e-41;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVVIVGRILSGRPVPIPREVLVYQEFDEMEECASHLPVTEQGMQLAEQFKQKALGLLOT 60
Db 528 SVVIVGRILSGRPVPIPREVLVYQEFDEMEECASHLPVTEQGMQLAEQFKQKALGLLOT 587
QY 61 ATKQAEAAAPVVEKWRAL 79
Db 588 ATKQAEAAAPVVEKWRAL 606
RESULT 10
AAR29850
ID AAR29850 standard; protein; 768 AA.
XX AC AAR29850;
XX XX
XX 25-MAR-2003 (revised)
DT

DT 26-APR-1993 (first entry)
XX DE
XX HCV NS2-NS4 peptide N23N15.
XX KW
XX Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum; HC;
KW transcriptase; cDNA; primer; allele.
XX OS
XX Hepatitis C virus.
XX
XX Key Location/Qualifiers
FH Misc-difference 19 /label= Pro, Leu
FT
FT Misc-difference 24 /label= His, Arg
FT
FT Misc-difference 60 /label= Thr, Ala
FT
FT Misc-difference 71 /label= Ser, Thr
FT
FT Misc-difference 88 /label= Lys, Arg
FT
FT Misc-difference 140 /label= Ile, Val
FT
FT Misc-difference 241 /label= Thr, Ile
FT
FT Misc-difference 274 /label= Val, Ile
FT
FT Misc-difference 381 /label= Pro, Ser
FT
FT Misc-difference 475 /label= Tyr, His
FT
FT Misc-difference 483 /label= Gln, Lys
FT
FT Misc-difference 485 /label= Thr, Ala
FT
FT Misc-difference 493 /label= Tyr, Phe
FT
FT Misc-difference 496 /label= Thr, Ala
FT
XX EP518313-A2.
XX PN
XX 16-DEC-1992.
XX PD
XX 11-JUN-1992; 92EP-00109812.
XX PF
XX 11-JUN-1991; 91JP-00139268.
XX PR
XX 12-JUL-1991; 91JP-00172794.
XX PR
XX 07-OCT-1991; 91JP-00287008.
XX PR
XX 16-DEC-1991; 91JP-00332329.
XX PR
XX 20-APR-1992; 92JP-00099957.
XX PR
XX (MITU) MITSUBISHI KASEI CORP.
XX PA
XX Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayaishi N;
XX PI
XX WPI; 1992-417213/51.
XX DR
XX N-PSDB; AAQ32481.
XX DR
XX New hepatitis C virus gene and its encoded protein - used for diagnosing
PT and vaccinating against hepatitis C virus infections.
XX PT
XX Disclosure; Page 134-38; 305pp; English.
XX PS
XX The sequences given in AAR29660, AAR29559-60 and AAR29843-51 were encoded
CC by clones which encode the NS2-NS4 regions of the Hepatitis C Virus (HCV)
CC gene of the invention. These sequences were isolated from the serum of a
CC patient suffering from hepatitis C (HC). The NS2-NS4 RNA sequences were
CC converted into cDNA using transcriptase in the presence of one of the
CC primer sequences given in AAQ32553-64. The cDNA sequences were then
CC amplified using primer pairs. The cDNA sequences isolated represent
CC different alleles of the same region of the HCV gene. Sequence
CC comparisons of these clones showed that it is possible for a patient to
CC carry more than one HCV strain at one time. See also AAQ32436. (Updated

```
CC on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 768 AA;

  Query Match      100.0%; Score 394; DB 2; Length 768;
  Best Local Similarity 100.0%; Pred. No. 9.1e-41;
  Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVVIVGRILSGRPAVIPDREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLIQT 60
Db 528 SVVIVGRILSGRPAVIPDREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLIQT 587

Qy 61 ATKQAEAAAPVVESKWRAL 79
Db 588 ATKQAEAAAPVVESKWRAL 606

RESULT 11
AAR29869
ID AAR29869 standard; protein; 768 AA.
XX
AC AAR29869;
XX
DT 25-MAR-2003 (revised)
DT 26-APR-1993 (first entry)
XX
DE HCV NS2-NS4 peptide N23N15B-1.
XX
KW Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum; HC;
KW transcriptase; cDNA; primer; allele.
XX
OS Hepatitis C virus.
XX
PN EP518313-A2.
XX
PD 16-DEC-1992.
XX
PF 11-JUN-1992; 92EP-00109812.
XX
PR 11-JUN-1991; 91JP-00139268.
PR 12-JUL-1991; 91JP-00172794.
PR 07-OCT-1991; 91JP-00287008.
PR 16-DEC-1991; 91JP-00332329.
PR 20-APR-1992; 92JP-00099957.
XX
PA (MITU ) MITSUBISHI KASEI CORP.
XX
PI Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;
XX
DR WPI; 1992-417213/51.
XX
DR N-PSDB; AAQ32500.
XX
PT New hepatitis C virus gene and its encoded protein - used for diagnosing
PT and vaccinating against hepatitis C virus infections.
XX
PS Disclosure; Page 182-86; 305pp; English.
XX
CC The sequences given in AAR29852-70 are encoded by various clones which
CC were used in the isolation of the NS2-NS4 regions of the Hepatitis C
CC Virus (HCV) gene of the invention (see also AAR29660, AAR29559-60 and
CC AAR29843-51). These RNA sequences were isolated from the serum of a
CC patient suffering from hepatitis C (HC). The isolated RNA sequences were
CC converted into cDNA using transcriptase in the presence of one of the
CC primer sequences given in AAQ32578-79. The sequences were then amplified
CC using primer pairs. The cDNA sequences isolated represent different
CC alleles of the same region of the HCV gene. Sequence comparisons of these
CC clones showed that it is possible for a patient to carry more than one
CC HCV strain at one time. See also AAQ32436. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 768 AA;

  Query Match      100.0%; Score 394; DB 2; Length 768;

Best Local Similarity 100.0%; Pred. No. 9.1e-41;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVVIVGRILSGRPAVIPDREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLIQT 60
Db 528 SVVIVGRILSGRPAVIPDREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLIQT 587

Qy 61 ATKQAEAAAPVVESKWRAL 79
Db 588 ATKQAEAAAPVVESKWRAL 606

RESULT 12
AAR29660
ID AAR29660 standard; protein; 1188 AA.
XX
AC AAR29660;
XX
DT 25-MAR-2003 (revised)
DT 26-APR-1993 (first entry)
XX
DE HCV NS2-NS4 peptide MX25N15.
XX
KW Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum; HC;
KW transcriptase; cDNA; primer; allele.
XX
OS Hepatitis C virus.
XX
Key Location/Qualifiers
FH Misc-difference 56
FT /label= Ala, Thr
FT Misc-difference 59
FT /label= Phe, Leu
FT Misc-difference 70
FT /label= Met, Leu
FT Misc-difference 74
FT /label= Ala, Pro
FT Misc-difference 82
FT /label= Glu, Asp
FT Misc-difference 110
FT /label= Lys, Arg
FT Misc-difference 174
FT /label= Gly, Ala
FT Misc-difference 175
FT /label= Leu, Ile
FT Misc-difference 177
FT /label= Gln, Arg
FT Misc-difference 178
FT /label= Met, Val
FT Misc-difference 190
FT /label= Met, Ala
FT Misc-difference 197
FT /label= Ala, Val
FT Misc-difference 207
FT /label= Leu, Phe
FT Misc-difference 208
FT /label= Met, Val
FT Misc-difference 218
FT /label= Val, Ile
FT Misc-difference 220
FT /label= Asp, Val
FT Misc-difference 223
FT /label= Thr, Ala
FT Misc-difference 227
FT /label= Asp, His
FT Misc-difference 244
FT /label= Ala, Val
FT Misc-difference 258
FT /label= Asp, Glu
FT Misc-difference 267
FT /label= Leu, Ser
FT Misc-difference 277
FT /label= Asn, Arg, Lys
```


FT Misc-difference 279 /label= Ile, Leu
 FT Misc-difference 281 /label= Leu, Phe
 FT Misc-difference 343 /label= Ile, Val
 FT Misc-difference 385 /label= Gly, Arg
 FT Misc-difference 439 /label= Leu, Pro
 FT Misc-difference 444 /label= His, Arg
 FT Misc-difference 480 /label= Thr, Ala
 FT Misc-difference 491 /label= Ser, Thr
 FT Misc-difference 508 /label= Lys, Arg
 FT Misc-difference 560 /label= Ile, Val
 FT Misc-difference 661 /label= Thr, Ile
 FT Misc-difference 694 /label= Val, Ile
 FT Misc-difference 801 /label= Pro, Ser
 FT Misc-difference 895 /label= Tyr, His
 FT Misc-difference 903 /label= Glu, Lys
 FT Misc-difference 905 /label= Thr, Ala
 FT Misc-difference 913 /label= Tyr, Phe
 FT Misc-difference 916 /label= Thr, Ala
 XX
 PN EP518313-A2.
 XX
 PD 16-DEC-1992.
 XX
 PF 11-JUN-1992; 92EP-00109812.
 XX
 PR 11-JUN-1991; 91JP-00139268.
 PR 12-JUL-1991; 91JP-00172794.
 PR 07-OCT-1991; 91JP-00287008.
 PR 16-DEC-1991; 91JP-00332329.
 PR 20-APR-1992; 92JP-00099957.
 XX
 PA (MITU) MITSUBISHI KASEI CORP.
 XX
 PI Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;
 XX
 DR WPI; 1992-417213/51.
 DR N-PSDB; AAQ32442.
 XX
 PT New hepatitis C virus gene and its encoded protein - used for diagnosing
 PT and vaccinating against hepatitis C virus infections.
 XX
 PS Disclosure; Page 139-45; 305pp; English.
 XX
 CC The sequences given in AAR29660, AAR29559-60 and AAR29843-51 were encoded
 CC by clones which encode the NS2-NS4 regions of the Hepatitis C Virus (HCV)
 CC gene of the invention. These sequences were isolated from the serum of a
 CC patient suffering from hepatitis C (HC). The NS2-NS4 RNA sequences were
 CC converted into cDNA using transcriptase in the presence of one of the
 CC primer sequences given in AAQ32553-64. The cDNA sequences were then
 CC amplified using primer pairs. The cDNA sequences isolated represent
 CC different alleles of the same region of the HCV gene. Sequence
 CC comparisons of these clones showed that it is possible for a patient to
 CC carry more than one HCV strain at one time. See also AAQ32436. (Updated
 CC on 25-MAR-2003 to correct PN field.)

SQ Sequence 1188 AA;
 Query Match 100.0%; Score 394; DB 2; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 1.6e-40;
 Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SVVIVGRIILSGRPAVIPDREVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLLOT 60
 Db 948 SVVIVGRIILSGRPAVIPDREVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLLOT 1007
 QY 61 ATKQAEAAAPVVEKWRAL 79
 Db 1008 ATKQAEAAAPVVEKWRAL 1026
 RESULT 13
 AAR29527
 ID AAR29527 standard; protein; 2510 AA.
 XX
 AC AAR29527;
 XX
 DT 25-MAR-2003 (revised)
 DT 26-APR-1993 (first entry)
 XX
 DE HCV antigen T7N1-30.
 XX
 KW Clone; Hepatitis C Virus; HCV; core-envelope; NS1(gp70); NS2-NS4;
 KW NS4-NS5; region; diagnostic method; antibody; suppress; control;
 KW proteolytic; process; precursor; polypeptide.
 XX
 OS Hepatitis C virus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 2212 /note= "Nonsense codon"
 XX
 PN EP518313-A2.
 XX
 PD 16-DEC-1992.
 XX
 PF 11-JUN-1992; 92EP-00109812.
 XX
 PR 11-JUN-1991; 91JP-00139268.
 PR 12-JUL-1991; 91JP-00172794.
 PR 07-OCT-1991; 91JP-00287008.
 PR 16-DEC-1991; 91JP-00332329.
 PR 20-APR-1992; 92JP-00099957.
 XX
 PA (MITU) MITSUBISHI KASEI CORP.
 XX
 PI Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;
 XX
 DR WPI; 1992-417213/51.
 DR N-PSDB; AAQ32436.
 XX
 PT New hepatitis C virus gene and its encoded protein - used for diagnosing
 PT and vaccinating against hepatitis C virus infections.
 XX
 PS Claim 1 and 3; Page 259-272; 305pp; English.
 XX
 CC This sequence was encoded by the Hepatitis C Virus (HCV) gene of the
 CC invention. The HCV gene is useful in the development of a diagnostic
 CC method which is more accurate and effective than conventional ones, in
 CC the detection of antibodies raised against a wide range of HCVs which
 CC have been hardly detected before. The complete gene may be used in an
 CC vitro screening system for a substance capable of specifically suppressing
 CC or controlling a proteolytic processing of a precursor polypeptide of
 CC HCV. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 2510 AA;
 Query Match 100.0%; Score 394; DB 2; Length 2510;
 Best Local Similarity 100.0%; Pred. No. 4.4e-40;

```
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVWIVGRIILSGRPAVDPREVLYQEFDEMEBCASHLPYIEQGMQLABQFKQKALGLIQT 60
    |||||
Db 1679 SVWIVGRIILSGRPAVDPREVLYQEFDEMEBCASHLPYIEQGMQLABQFKQKALGLIQT 1738
QY 61 ATKQAEAAAAPVVEKWRAL 79
    |||||
Db 1739 ATKQAEAAAAPVVEKWRAL 1757

RESULT 14
AAR68864
ID AAR68864 standard; protein; 3010 AA.
XX
AC AAR68864;
XX
DT 06-DEC-1995 (first entry)
XX
DE Hepatitis C virus RNA helicase.
XX
KW Hepatitis C virus; HCV; non-A non-B; helicase gene; RNA helicase;
KW baculovirus; recombinant production.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT 196..198
FT Region /label= N-linked glycosylation site
FT 209..211
FT Region /label= N-linked glycosylation site
FT 234..236
FT Region /label= N-linked glycosylation site
FT 250..252
FT Region /label= N-linked glycosylation site
FT 305..307
FT Region /label= N-linked glycosylation site
FT 325..327
FT Region /label= N-linked glycosylation site
FT 417..419
FT Region /label= N-linked glycosylation site
FT 423..425
FT Region /label= N-linked glycosylation site
FT 430..432
FT Region /label= N-linked glycosylation site
FT 448..450
FT Region /label= N-linked glycosylation site
FT 532..534
FT Region /label= N-linked glycosylation site
FT 556..558
FT Region /label= N-linked glycosylation site
FT 576..578
FT Region /label= N-linked glycosylation site
FT 623..625
FT Region /label= N-linked glycosylation site
FT 645..647
FT Region /label= N-linked glycosylation site
FT 1213..1215
FT Region /label= N-linked glycosylation site
FT 1255..1257
FT Region /label= N-linked glycosylation site
FT 2041..2043
FT Region /label= N-linked glycosylation site
FT 2077..2079
FT Region /label= N-linked glycosylation site
FT 2240..2242
FT Region /label= N-linked glycosylation site
FT 2788..2790
FT Region /label= N-linked glycosylation site
XX
PN JP06319583-A.
XX
PD 22-NOV-1994.
```

```
XX 18-SEP-1992; 92JJP-00249241.
PF
XX 18-SEP-1992; 92JJP-00249241.
PR
XX (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.
PA
XX WPI; 1995-040330/06.
DR N-PSDB; AAO81559.
DR
XX of hepatitis C virus helicase gene in baculovirus - useful for large
XX scale prodn. of RNA helicase.
XX
XX Claim 1; Fig 1-4; 9pp; Japanese.
XX
XX AAQ81559 encodes AAR68864 hepatitis C virus (HCV) RNA helicase. The DNA
XX was used in the construction of an expression vector, which was used to
XX transform a baculovirus host. The transformed baculovirus could then be
XX used for the recombinant prodn. of HCV RNA helicase
XX
XX Sequence 3010 AA;
SQ
Query Match 100.0%; Score 394; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 5.6e-40;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVWIVGRIILSGRPAVDPREVLYQEFDEMEBCASHLPYIEQGMQLABQFKQKALGLIQT 60
    |||||
Db 1679 SVWIVGRIILSGRPAVDPREVLYQEFDEMEBCASHLPYIEQGMQLABQFKQKALGLIQT 1738
QY 61 ATKQAEAAAAPVVEKWRAL 79
    |||||
Db 1739 ATKQAEAAAAPVVEKWRAL 1757

RESULT 15
AAR82694
ID AAR82694 standard; protein; 3010 AA.
XX
AC AAR82694;
XX
DT 16-OCT-2003 (revised)
DT 14-NOV-1996 (first entry)
XX
XX Partial HCV non-structural polyprotein.
DE
XX
KW proteinase; hepatitis C virus; screening; inhibitor; proteolytic;
KW identification; cleavage.
XX
OS Hepatitis C virus; Virus.
XX
XX Key Location/Qualifiers
FH 898..1233
FT Protein /note= "partial proteinase; see AAR82692"
FT Protein 992..1907
FT /note= "partial proteinase; see AAR82693"
XX
XX JP07184648-A.
XX
XX 25-JUL-1995.
XX
XX 05-FEB-1993; 93JJP-00018854.
XX
XX 07-FEB-1992; 92JJP-00022657.
PR 18-SEP-1992; 92JJP-00249240.
PR 04-DEC-1992; 92JJP-00325303.
XX
XX (KAEN/) KAENNO K.
PA (SUMQ) SUMITOMO METAL IND LTD.
PA (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.
XX
XX WPI; 1995-287962/38.
DR N-PSDB; AAT03960.
DR
```

XX An HCV proteinase active substance - which has activity as an anti-HCV
PT agent and can be used to screen for proteinase inhibitors.
XX Disclosure; Page 39-48; 52pp; Japanese.
PS
XX The present sequence is a partial Hepatitis C Virus (HCV) polyprotein
CC from the non-structural region. Partial proteinase sequences (AAR82692-
CC 93) are contained within this sequence. The proteinases can be used as
CC anti-HCV agents. They can also be used to screen cpds. for their ability
CC to inhibit their proteolytic activity. In this way, proteinase inhibitors
CC can be identified. (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 3010 AA;
Query Match 100.0%; Score 394; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 5.6e-40;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVVIVGRILSGRPAVIPDREVLYQEFDEMESCASHLPYIEQGMQLAEQFKKALGLLOT 60
Db 1679 SVVIVGRILSGRPAVIPDREVLYQEFDEMESCASHLPYIEQGMQLAEQFKKALGLLOT 1738
QY 61 ATKQAEAAAPVVESKWRAL 79
Db 1739 ATKQAEAAAPVVESKWRAL 1757
Search completed: August 12, 2005, 14:14:59
Job time : 83.8821 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:04:54 ; Search time 22.1278 Seconds
(without alignments)
266.510 Million cell updates/sec

Title: US-09-758-308-3

Perfect score: 394

Sequence: 1 SWVIVGRILSGRPAVIPDR.....TATKQAEAAAPVSEKMRAL 79

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	389	98.7	247	1 US-08-324-977-44	Sequence 44, Appl
2	389	98.7	247	2 US-08-384-616-44	Sequence 44, Appl
3	389	98.7	247	2 US-08-904-686A-44	Sequence 44, Appl
4	389	98.7	247	3 US-09-315-850-44	Sequence 44, Appl
5	389	98.7	1692	3 US-09-263-933-4	Sequence 4, Appl
6	389	98.7	1692	3 US-09-263-933-11	Sequence 11, Appl
7	389	98.7	1692	3 US-09-263-933-18	Sequence 18, Appl
8	389	98.7	1692	4 US-09-919-901-4	Sequence 4, Appl
9	389	98.7	1692	4 US-09-919-901-11	Sequence 11, Appl
10	389	98.7	1692	4 US-09-919-901-18	Sequence 18, Appl
11	389	98.7	1692	4 US-10-191-966-4	Sequence 4, Appl
12	389	98.7	1692	4 US-10-191-966-11	Sequence 11, Appl
13	389	98.7	1692	4 US-10-191-966-18	Sequence 18, Appl
14	389	98.7	2013	1 US-08-324-977-12	Sequence 12, Appl
15	389	98.7	2013	2 US-08-384-616-12	Sequence 12, Appl
16	389	98.7	2013	2 US-08-904-686A-12	Sequence 12, Appl
17	389	98.7	2013	3 US-09-315-850-12	Sequence 12, Appl
18	389	98.7	2201	3 US-08-952-981A-2	Sequence 2, Appl
19	389	98.7	2307	3 US-09-263-933-2	Sequence 2, Appl
20	389	98.7	2307	3 US-09-263-933-9	Sequence 9, Appl
21	389	98.7	2307	3 US-09-263-933-16	Sequence 16, Appl
22	389	98.7	2307	4 US-09-919-901-2	Sequence 2, Appl
23	389	98.7	2307	4 US-09-919-901-9	Sequence 9, Appl
24	389	98.7	2307	4 US-09-919-901-16	Sequence 16, Appl
25	389	98.7	2307	4 US-10-191-966-2	Sequence 2, Appl
26	389	98.7	2307	4 US-10-191-966-9	Sequence 9, Appl
27	389	98.7	2307	4 US-10-191-966-16	Sequence 16, Appl

28	389	98.7	2620	1 US-08-324-977-32	Sequence 32, Appl
29	389	98.7	2620	2 US-08-384-616-32	Sequence 32, Appl
30	389	98.7	2620	2 US-08-904-686A-32	Sequence 32, Appl
31	389	98.7	2620	3 US-09-315-850-32	Sequence 32, Appl
32	389	98.7	2621	1 US-08-324-977-36	Sequence 36, Appl
33	389	98.7	2621	2 US-08-384-616-36	Sequence 36, Appl
34	389	98.7	2621	2 US-08-904-686A-36	Sequence 36, Appl
35	389	98.7	2621	3 US-09-315-850-36	Sequence 36, Appl
36	389	98.7	3010	1 US-08-324-977-2	Sequence 2, Appl
37	389	98.7	3010	1 US-08-324-977-14	Sequence 14, Appl
38	389	98.7	3010	2 US-08-384-616-2	Sequence 2, Appl
39	389	98.7	3010	2 US-08-384-616-14	Sequence 14, Appl
40	389	98.7	3010	2 US-08-904-686A-2	Sequence 2, Appl
41	389	98.7	3010	2 US-08-904-686A-14	Sequence 14, Appl
42	389	98.7	3010	3 US-09-315-850-2	Sequence 2, Appl
43	389	98.7	3010	3 US-09-315-850-14	Sequence 14, Appl
44	385	97.7	87	1 US-08-685-764-2	Sequence 2, Appl
45	382	97.0	3010	3 US-09-014-416-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-324-977-44
; Sequence 44, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McIeland &
; ADDRESSER: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281

; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-324-977-44

Query Match 98.7%; Score 389; DB 1; Length 247;
Best Local Similarity 96.2%; Pred. No. 1.4e-42;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVVIVGRILSGRPAIVDPREVLVYQEFDEMEECASHLPYIEQGMQLARQFKKALGLLQ 60
DB 64 SVVIVGRILSGRPAIVDPRELLYQEFDEMEECASHLPYIEQGMQLARQFKKALGLLQ 123
QY 61 ATKQAEAAAAPVVESKWRAL 79
DB 124 ATKQAEAAAAPVVESKWRAL 142

RESULT 2

US-08-384-616-44
; Sequence 44, Application US/08384616
; Patent No. 5847101

; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeand &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281

; REFERENCE/DOCKET NUMBER: 900703B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-384-616-44

Query Match 98.7%; Score 389; DB 2; Length 247;
Best Local Similarity 96.2%; Pred. No. 1.4e-42;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVVIVGRILSGRPAIVDPREVLVYQEFDEMEECASHLPYIEQGMQLARQFKKALGLLQ 60
DB 64 SVVIVGRILSGRPAIVDPRELLYQEFDEMEECASHLPYIEQGMQLARQFKKALGLLQ 123
QY 61 ATKQAEAAAAPVVESKWRAL 79
DB 124 ATKQAEAAAAPVVESKWRAL 142

RESULT 3

US-08-904-686A-44
; Sequence 44, Application US/08904686A
; Patent No. 5998130

; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeand &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/635,451
;; FILING DATE: 28-DEC-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McLeLand, Le-Nhung
;; REGISTRATION NUMBER: 31,541
;; REFERENCE/DOCKET NUMBER: 900703G
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 659-2930
;; TELEFAX: (202) 887-0357
;; INFORMATION FOR SEQ ID NO: 44:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 247 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-904-686A-44

Query Match 98.7%; Score 389; DB 2; Length 247;
Best Local Similarity 96.2%; Pred. No. 1.4e-42;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVVIVGRIILSGRPAVIDPDEVLYQEFDEMEECASHLPYIEQGMQLAEQFKOKALGLLOT 60
Db 64 SVVIVGRIILSGRPAVIDPDEVLYQEFDEMEECASHLPYIEQGMQLAEQFKOKALGLLOT 123

QY 61 ATKQAEAAAAPVVESKWRAL 79
Db 124 ATKQAEAAAAPVVESKWRAL 142

RESULT 4
US-09-315-850-44
; Sequence 44, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLand &
; ADDRESSER: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,850
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686
; FILING DATE: 01-AUG-1997
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605

;; FILING DATE: 09-NOV-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/099,706
;; FILING DATE: 30-JUL-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/769,996
;; FILING DATE: 02-OCT-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/635,451
;; FILING DATE: 28-DEC-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McLeLand, Le-Nhung
;; REGISTRATION NUMBER: 31,541
;; REFERENCE/DOCKET NUMBER: 900703G
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 659-2930
;; TELEFAX: (202) 887-0357
;; INFORMATION FOR SEQ ID NO: 44:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 247 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-09-315-850-44

Query Match 98.7%; Score 389; DB 3; Length 247;
Best Local Similarity 96.2%; Pred. No. 1.4e-42;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVVIVGRIILSGRPAVIDPDEVLYQEFDEMEECASHLPYIEQGMQLAEQFKOKALGLLOT 60
Db 64 SVVIVGRIILSGRPAVIDPDEVLYQEFDEMEECASHLPYIEQGMQLAEQFKOKALGLLOT 123

QY 61 ATKQAEAAAAPVVESKWRAL 79
Db 124 ATKQAEAAAAPVVESKWRAL 142

RESULT 5
US-09-263-933-4
; Sequence 4, Application US/09263933
; Patent No. 6280940
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/263,933
; CURRENT FILING DATE: 1999-03-08
; EARLIER APPLICATION NUMBER: 09/129,611
; EARLIER FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-263-933-4

Query Match 98.7%; Score 389; DB 3; Length 1692;
Best Local Similarity 96.2%; Pred. No. 1.8e-41;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVVIVGRIILSGRPAVIDPDEVLYQEFDEMEECASHLPYIEQGMQLAEQFKOKALGLLOT 60
Db 958 SVVIVGRIILSGRPAVIDPDEVLYQEFDEMEECASHLPYIEQGMQLAEQFKOKALGLLOT 1017

QY 61 ATKQAEAAAAPVVESKWRAL 79
Db 1018 ATKQAEAAAAPVVESKWRAL 1036

```
RESULT 6
US-09-263-933-11
; Sequence 11, Application US/09263933
; Patent No. 6280940
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/263,933
; CURRENT FILING DATE: 1999-03-08
; EARLIER APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION:
US-09-263-933-11

Query Match      98.7%; Score 389; DB 3; Length 1692;
Best Local Similarity 96.2%; Pred. No. 1.8e-41;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVWIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLAEQFKQKALGLLOT 60
Db 958 SVWIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLAEQFKQKALGLLOT 1017

Qy 61 ATKQAEAAAAPVVESKWRAL 79
Db 1018 ATKQAEAAAAPVVESKWRAL 1036

RESULT 7
US-09-263-933-18
; Sequence 18, Application US/09263933
; Patent No. 6280940
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/263,933
; CURRENT FILING DATE: 1999-03-08
; EARLIER APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION:
US-09-263-933-18

Query Match      98.7%; Score 389; DB 3; Length 1692;
Best Local Similarity 96.2%; Pred. No. 1.8e-41;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVWIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLAEQFKQKALGLLOT 60
Db 958 SVWIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLAEQFKQKALGLLOT 1017

Qy 61 ATKQAEAAAAPVVESKWRAL 79
Db 1018 ATKQAEAAAAPVVESKWRAL 1036

RESULT 8
US-09-919-901-4
; Sequence 4, Application US/09919901
; Patent No. 6599738
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/919,901
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/263,933
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION:
US-09-919-901-4

Query Match      98.7%; Score 389; DB 4; Length 1692;
Best Local Similarity 96.2%; Pred. No. 1.8e-41;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVWIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLAEQFKQKALGLLOT 60
Db 958 SVWIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLAEQFKQKALGLLOT 1017

Qy 61 ATKQAEAAAAPVVESKWRAL 79
Db 1018 ATKQAEAAAAPVVESKWRAL 1036

RESULT 9
US-09-919-901-11
; Sequence 11, Application US/09919901
; Patent No. 6599738
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/919,901
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/263,933
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION:
US-09-919-901-11

Query Match      98.7%; Score 389; DB 4; Length 1692;
Best Local Similarity 96.2%; Pred. No. 1.8e-41;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVWIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLAEQFKQKALGLLOT 60
```

```
Qy 1 SVWIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLAEQFKQKALGLLOT 60
```



```
Db 958 SVVIVGRIILSGRPAIVDPRELYQEFDEMEECASHLPYIEQGMQLAEQFKOKALGLLOT 1017
Qy 61 ATKQAEAAAPVVESKWRAL 79
Db 1018 ATKQAEAAAPVVESKWRAL 1036

RESULT 10
US-09-919-901-18
; Sequence 4, Application US/09919901
; Patent No. 6599738
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/919,901
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/263,933
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION:
US-09-919-901-18

Query Match 98.7%; Score 389; DB 4; Length 1692;
Best Local Similarity 96.2%; Pred. No. 1.8e-41;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVVIVGRIILSGRPAIVDPRELYQEFDEMEECASHLPYIEQGMQLAEQFKOKALGLLOT 60
Db 958 SVVIVGRIILSGRPAIVDPRELYQEFDEMEECASHLPYIEQGMQLAEQFKOKALGLLOT 1017

RESULT 11
US-10-191-966-4
; Sequence 4, Application US/10191966
; Patent No. 6790612
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/10/191,966
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US/09/263,933
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION:
US-10-191-966-4
```

```
Query Match 98.7%; Score 389; DB 4; Length 1692;
Best Local Similarity 96.2%; Pred. No. 1.8e-41;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVVIVGRIILSGRPAIVDPRELYQEFDEMEECASHLPYIEQGMQLAEQFKOKALGLLOT 60
Db 958 SVVIVGRIILSGRPAIVDPRELYQEFDEMEECASHLPYIEQGMQLAEQFKOKALGLLOT 1017

RESULT 12
US-10-191-966-11
; Sequence 11, Application US/10191966
; Patent No. 6790612
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/10/191,966
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US/09/263,933
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION:
US-10-191-966-11

Query Match 98.7%; Score 389; DB 4; Length 1692;
Best Local Similarity 96.2%; Pred. No. 1.8e-41;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVVIVGRIILSGRPAIVDPRELYQEFDEMEECASHLPYIEQGMQLAEQFKOKALGLLOT 60
Db 958 SVVIVGRIILSGRPAIVDPRELYQEFDEMEECASHLPYIEQGMQLAEQFKOKALGLLOT 1017

RESULT 13
US-10-191-966-18
; Sequence 18, Application US/10191966
; Patent No. 6790612
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/10/191,966
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US/09/263,933
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; OTHER INFORMATION:
US-10-191-966-18
```

```
; SEQ ID NO 18
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION:
US-10-191-966-18

Query Match          98.7%; Score 389; DB 4; Length 1692;
Best Local Similarity 96.2%; Pred. No. 1.8e-41;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVVIGRIILGRPAIVDPREVLVYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLIQT 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 958 SVVIGRIILGRPAIVDPRELLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLIQT 1017
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 61 ATKQAEAAAAPVVESKWRAL 79
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1018 ATKQAEAAAAPVVESKWRAL 1036
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 14
US-08-324-977-12
; Sequence 12, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAWIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLend &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:

;
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2013 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-324-977-12

Query Match          98.7%; Score 389; DB 1; Length 2013;
Best Local Similarity 96.2%; Pred. No. 2.2e-41;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVVIGRIILGRPAIVDPREVLVYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLIQT 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1679 SVVIGRIILGRPAIVDPRELLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLIQT 1738
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 61 ATKQAEAAAAPVVESKWRAL 79
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1739 ATKQAEAAAAPVVESKWRAL 1757
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 15
US-08-384-616-12
; Sequence 12, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAWIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLend &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703B
; TELECOMMUNICATION INFORMATION:
```

; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2013 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-384-616-12

Query Match 98.7%; Score 389; DB 2; Length 2013;
Best Local Similarity 96.2%; Pred. No. 2.2e-41;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVVIVGRILSGRPVIVDPREVLVYQEFDEMEECASHLPVIEQGMQLAQPKKALGLLQT 60
Db 1679 SVVIVGRILSGRPVIVDPRELLVYQEFDEMEECASHLPVIEQGMQLAQPKKALGLLQT 1738

QY 61 ATKQAEAAAPVVESKWRAL 79
Db 1739 ATKQAEAAAPVVESKWRAL 1757

Search completed: August 12, 2005, 14:24:48
Job time : 22.1278 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:04:14 ; Search time 15.7224 Seconds
(without alignments)
483.460 Million cell updates/sec

Title: US-09-758-308-3

Perfect score: 394

Sequence: 1 SWVIGRILSGRPVAVIPDR.....TATKQAEAAAPVVEKWRAL 79

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	394	100.0	3010	1 GNVVCJ	genome polyprotein
2	389	98.7	3010	1 GNVVC	genome polyprotein
3	386	98.0	3010	1 A45573	genome polyprotein
4	386	98.0	3010	1 GNVVTW	genome polyprotein
5	385	97.7	3010	1 S18030	genome polyprotein
6	333	84.5	492	2 PS0326	polyprotein - hepa
7	332	84.3	3011	1 S40770	genome polyprotein
8	328	83.2	3011	1 GNVVC3	genome polyprotein
9	324	82.2	3011	1 GNVVCH	genome polyprotein
10	320	81.2	716	2 JQ1366	polyprotein - hepa
11	254	64.5	3014	1 JCS620	genome polyprotein
12	247	62.7	142	2 PC1307	genome polyprotein
13	247	62.7	209	2 PC1306	genome polyprotein
14	229	58.1	876	2 PC2219	polypeptide - hepa
15	208	52.8	3033	1 GNVVJ8	genome polyprotein
16	205	52.0	3033	1 JQ1303	genome polyprotein
17	204	51.8	125	2 S35629	hypothetical prote
18	70.5	17.9	401	2 A89916	aspartokinase II (
19	67.5	17.1	980	2 S71090	peroxisome biogene
20	64.5	16.4	286	2 G71899	hypothetical prote
21	64.5	16.4	483	2 E97255	transcription term
22	64	16.2	144	1 C70437	ATP synthase F0 su
23	64	16.2	1168	2 S76195	hypothetical prote
24	64	16.2	1697	2 T00079	hypothetical prote
25	63.5	16.1	278	2 T44412	ABC transporter (A
26	63.5	16.1	423	2 AB1393	transcription term
27	63.5	16.1	423	2 AB1769	transcription term
28	63.5	16.1	446	2 A40896	Ca2+/calmodulin-de
29	63.5	16.1	1116	2 A84046	DNA polymerase III

30 63 16.0 433 2 D82077 methyl-accepting c
31 63 16.0 688 2 T40832 conserved hypother
32 63 16.0 808 2 A96791 hypothetical prote
33 63 16.0 3491 2 T43231 probable 6-deoxyar
34 62.5 15.9 268 2 AH2521 hypothetical prote
35 62.5 15.9 323 2 T19426 hypothetical prote
36 62.5 15.9 380 2 S31045 hypothetical prote
37 62.5 15.9 407 2 T16238 hypothetical prote
38 62.5 15.9 476 2 D71264 conserved hypother
39 62 15.7 485 2 G71527 probable pyruvate
40 62 15.7 911 2 JG7186 alpha-actinin-4 -
41 62 15.7 2777 2 D96746 hypothetical prote
42 61.5 15.6 215 2 A11773 hypothetical prote
43 61.5 15.6 828 2 AD0412 ATP-dependent heli
44 61.5 15.6 856 2 E75292 GDEF family prote
45 61.5 15.6 908 2 B69435 signal-transducing

ALIGNMENTS

RESULT 1

GNVVCJ

genome polyprotein - hepatitis C virus (strain J)

N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

C;Species: hepatitis C virus

C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004

C;Accession: A39253; PS0086

R;Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugamura, T.; Shimor

Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990

A;Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients

A;Reference number: A39253; MUID:91088550; PMID:2175903

A;Accession: A39253

A;Molecule type: genomic RNA

A;Residues: 1-3010 <KAT>

A;Cross-references: UNIPROT:P26662; GB:D90208; NID:G221610; PID:G221611

R;Kato, N.; Ohkoshi, S.; Shimotohno, K.

Proc. Jpn. Acad. 65B, 219-223, 1989

A;Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence vari

A;Reference number: PS0085

A;Accession: PS0086

A;Molecule type: genomic RNA

A;Residues: 2650-2707 <KA2>

A;Experimental source: Japanese isolate

C;Comment: The cleavage sites of this polyprotein have not been determined.

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; seri

F;2-115/Product: capsid protein C #status predicted <CPC>

F;116-191/Product: envelope protein M #status predicted <EPM>

F;192-389/Product: major envelope protein E #status predicted <MEB>

F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>

F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>

F;1007-1615/Product: hepatitis C virus genome polyprotein NS3 #status predicted <NS3>

F;1230-1237/Region: nucleotide-binding motif A (P-loop)

F;1312-1317/Region: nucleotide-binding motif B

F;1316-1319/Region: DEXH motif

F;1616-1862/Product: nonstructural protein NS4a #status predicted <NS4a>

F;1863-2013/Product: nonstructural protein NS4b #status predicted <NS4b>

F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

F;196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,...

Query Match 100.0%; Score 394; DB 1; Length 3010;

Best Local Similarity 100.0%; Pred. No. 9.1e-34;

Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SWVIGRILSGRPVAVIPDRVLYQFDEMEECASHLPYIEQGMOLAEQFKKALGLLOT 60

Db 1679 SWVIGRILSGRPVAVIPDRVLYQFDEMEECASHLPYIEQGMOLAEQFKKALGLLOT 1738

QY 61 ATKQAEAAAPVVEKWRAL 79

Db 1739 ATKQAEAAAPVVEKWRAL 1757

```
RESULT 2
GNWVTW
genome polyprotein - hepatitis C virus
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome isolated from human
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: A38465
R:Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.;
J. Virol. 65, 1105-1113, 1991
A:Title: Structure and organization of the hepatitis C virus genome isolated from human
A:Reference number: A38465; MUID:91140698; PMID:1847440
A:Accession: A38465
A:Molecule type: genomic RNA
A:Residues: 1-3010 <TAK>
A:Cross-references: UNIPROT:P26663; EMBL:M58335; NID:g329770; PIDN:AAA72945.1; PID:g329770
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <BPM>
F:192-389/Product: major envelope protein E #status predicted <MEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polyprotein NS5 #status predicted <NS5>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,22
Query Match 98.7%; Score 389; DB 1; Length 3010;
Best Local Similarity 96.2%; Pred. No. 3.1e-33;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVVIVGRILSGRPAIVDPREVLVYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLQ 60
Db 1679 SVVIVGRILSGRPAIVDPREVLVYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLQ 1738
QY 61 ATKQAEAAAPVVESKWRAL 79
Db 1739 ATKQAEAAAPVVESKWRAL 1757
RESULT 3
A45573
genome polyprotein - hepatitis C virus (strain JT)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome isolated from human
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: A45573
R:Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsubo, Y.; Cho, M.J.; Nakazawa, T.; Hijikata,
Virus Res. 23, 39-53, 1992
A:Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: B
A:Reference number: A45573; MUID:92295714; PMID:1318627
A:Accession: A45573
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3010 <TAN>
A:Cross-references: UNIPROT:Q00269; GB:D01168; NID:g221612; PIDN:BA01943.1;
A:Experimental source: HCV-JT
A:Note: sequence extracted from NCBI backbone (NCBI:106206, NCBI:P106207)
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <BPM>
F:192-389/Product: major envelope protein E #status predicted <MEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
```

```
F:1007-1615/Product: hepatitis C virus genome polyprotein NS5 #status predicted <NS5>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
Query Match 98.0%; Score 386; DB 1; Length 3010;
Best Local Similarity 96.2%; Pred. No. 6.6e-33;
Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVVIVGRILSGRPAIVDPREVLVYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLQ 60
Db 1679 SVVIVGRILSGRPAIVDPREVLVYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLQ 1738
QY 61 ATKQAEAAAPVVESKWRAL 79
Db 1739 ATKQAEAAAPVVESKWRAL 1757
RESULT 4
GNWVTW
genome polyprotein - hepatitis C virus (strain Taiwan)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome isolated from human
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: A40244
R:Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
A:Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the
A:Reference number: A40244; MUID:92230206; PMID:1314449
A:Accession: A40244
A:Molecule type: genomic RNA
A:Residues: 1-3010 <CHE>
A:Cross-references: UNIPROT:P29846; GB:M84754
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural
F:1-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <BPM>
F:192-389/Product: major envelope protein E #status predicted <MEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polyprotein NS5 #status predicted <NS5>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
Query Match 98.0%; Score 386; DB 1; Length 3010;
Best Local Similarity 96.2%; Pred. No. 6.6e-33;
Matches 76; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 SVVIVGRILSGRPAIVDPREVLVYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLQ 60
Db 1679 SVVIVGRILSGRPAIVDPREVLVYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLQ 1738
QY 61 ATKQAEAAAPVVESKWRAL 79
Db 1739 ATKQAEAAAPVVESKWRAL 1757
RESULT 5
S18030
genome polyprotein - hepatitis C virus (isolate JK1)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome isolated from human
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
```

A; Variety: isolate JKI
 C; Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
 C; Accession: S18030; S33570; A48332; S18029
 R; Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
 Submitted to the EMBL Data Library, September 1991
 A; Description: A whole genome of hepatitis C virus cDNA was isolated from a single patient
 A; Reference number: S18028
 A; Accession: S18030
 A; Molecule type: genomic RNA
 A; Residues: 1-3010 <HON>
 A; Cross-references: UNIPROT:Q68949; EMBL:X61596; NID:G59478; PIDN:CAA43793.1; PID:G59479
 A; Experimental source: isolate JKI from an individual
 R; Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
 Arch. Virol. 128, 163-169, 1993
 A; Title: Sequence analysis of putative structural regions of hepatitis C virus isolated
 A; Reference number: A48332; MUID:93119270; PMID:8380322
 A; Accession: S33570
 A; Molecule type: genomic RNA
 A; Residues: 1-547, 'T', 549-621, 'V', 623-624, 'S', 626-652, 'DL', 655-761, 'T', 763-782 <HOW>
 A; Cross-references: EMBL:X61591
 A; Note: this sequence is inconsistent with the nucleotide translation
 A; Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 320 as Trp, and TTC for residue 771 as Ser
 A; Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBIP:121748)
 C; Superfamily: hepatitis C virus genome polyprotein
 C; Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
 F; 2-115/Product: capsid protein C #status predicted <CPC>
 F; 116-191/Product: envelope protein M #status predicted <EPM>
 F; 192-389/Product: major envelope protein E #status predicted <MEE>
 F; 390-729/Product: nonstructural protein NS1 #status predicted <NS1>
 F; 730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
 F; 1007-1615/Product: hepatitis C virus polyprotein NS3 #status predicted <NS3>
 F; 1230-1237/Region: nucleotide-binding motif A (P-loop)
 F; 1312-1317/Region: nucleotide-binding motif B
 F; 1316-1319/Region: DEXH motif
 F; 1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
 F; 1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
 F; 2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
 F; 196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate (Asn)
 Query Match 97.7%; Score 385; DB 1; Length 3010;
 Best Local Similarity 96.2%; Pred. No. 8.5e-33;
 Matches 76; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SWVIGRIILSGRPAVIDPREVLYOEFDEMEECSHLPYIEQGMOLAEQFKKALGLLQTA 60
 DB 1679 SWVIGRIILSGRPAVIDPREVLYOEFDEMEECSHLPYIEQGMOLAEQFKKALGLLQTA 1738
 QY 61 ATKQAEAAAPVVEKWRAL 79
 DB 1739 ASKQAEAAAPVVEKWRAL 1757
 Query Match 84.3%; Score 332; DB 1; Length 3011;
 Best Local Similarity 79.5%; Pred. No. 4.3e-27;
 Matches 62; Conservative 9; Mismatches 7; Indels 0; Gaps 0;
 RESULT 6
 PS0326
 polyprotein - hepatitis C virus (isolate Fla) (fragments)
 C; Species: hepatitis C virus
 C; Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
 C; Accession: PS0326
 R; Li, J.S.; Tong, S.P.; Vitvitski, L.; Lepot, D.; Treppe, C.
 Gene 105, 167-172, 1991
 A; Title: Two French genotypes of hepatitis C virus: homology of the predominant genotype
 A; Reference number: PS0326; MUID:92039028; PMID:1718820
 A; Accession: PS0326
 A; Molecule type: genomic RNA
 A; Residues: 1-492 <LIJ>
 A; Cross-references: UNIPROT:Q91FE5; UNIPROT:Q36579; UNIPROT:Q36610; UNIPROT:Q03463; UNIE
 M60220
 A; Note: this sequence corresponds to nonstructural protein NS3 region
 A; Note: translation of the nucleotide sequence is not complete
 C; Superfamily: hepatitis C virus genome polyprotein
 C; Keywords: polyprotein

Query Match 84.5%; Score 333; DB 2; Length 492;
 Best Local Similarity 79.5%; Pred. No. 4.7e-28;
 Matches 62; Conservative 9; Mismatches 7; Indels 0; Gaps 0;
 QY 2 WVIVGRIILSGRPAVIDPREVLYOEFDEMEECSHLPYIEQGMOLAEQFKKALGLLQTA 61
 DB 231 WVIVGRIILSGRPAVIDPREVLYOEFDEMEECSHLPYIEQGMOLAEQFKKALGLLQTA 290
 QY 62 TKQAEAAAPVVEKWRAL 79
 DB 291 SRQAEIATPAVTNNQRL 308
 Query Match 84.5%; Score 332; DB 1; Length 3011;
 Best Local Similarity 79.5%; Pred. No. 4.3e-27;
 Matches 62; Conservative 9; Mismatches 7; Indels 0; Gaps 0;
 QY 2 WVIVGRIILSGRPAVIDPREVLYOEFDEMEECSHLPYIEQGMOLAEQFKKALGLLQTA 61
 DB 1680 WVIVGRIILSGRPAVIDPREVLYOEFDEMEECSHLPYIEQGMOLAEQFKKALGLLQTA 1739
 QY 62 TKQAEAAAPVVEKWRAL 79
 DB 1740 SRQAEIATPAVTNNQRL 1757
 Query Match 84.3%; Score 332; DB 1; Length 3011;
 Best Local Similarity 79.5%; Pred. No. 4.3e-27;
 Matches 62; Conservative 9; Mismatches 7; Indels 0; Gaps 0;
 RESULT 8
 GNMVC3
 genome polyprotein - hepatitis C virus (strain HCV-1)
 N; Contains: capsid protein C; envelope protein M; hepatitis C virus polyprotein NS3 region
 A; Note: translation of the nucleotide sequence is not complete
 C; Superfamily: hepatitis C virus genome polyprotein
 C; Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
 C; Accession: A39166; PQ0403; PQ0404

genome polyprotein - hepatitis C virus
 N; Contains: capsid protein C; envelope protein M; hepatitis C virus polyprotein NS3 region
 A; Note: translation of the nucleotide sequence is not complete
 C; Superfamily: hepatitis C virus genome polyprotein
 C; Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
 C; Accession: A39166; PQ0403; PQ0404

genome polyprotein - hepatitis C virus (strain HCV-1)
 N; Contains: capsid protein C; envelope protein M; hepatitis C virus polyprotein NS3 region
 A; Note: translation of the nucleotide sequence is not complete
 C; Superfamily: hepatitis C virus genome polyprotein
 C; Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
 C; Accession: A39166; PQ0403; PQ0404

R;Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Coi Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
A;Title: Genetic organization and diversity of the hepatitis C virus.
A;Reference number: A39166; MUID:91172826; PMID:1848704
A;Accession: A39166
A;Molecule type: mRNA
A;Residues: 1-3011 <CHO>
A;Cross-references: UNIPROT:P26664; GB:M62321; NID:g329873; PIDN:AAA45676.1; PID:g329874
R;Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Feuthner, J.F.; Follett, E.; Yap, P.L J. Gen. Virol. 73, 1131-1141, 1992
A;Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e
A;Reference number: PQ0393; MUID:92268871; PMID:1316939
A;Accession: PQ0403
A;Molecule type: genomic RNA
A;Residues: 1577-1633 <CHA>
A;Cross-references: DBJ:D10128
A;Experimental source: isolates E-b16
A;Accession: PQ0404
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1577-1633 <CH2>
A;Experimental source: isolates E-b17
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
F;1-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;730-1006/Product: nonstructural protein NS1 #status predicted <NS1>
F;1007-1615/Product: nonstructural protein NS2 #status predicted <NS2>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F;196, 209, 234, 305, 325, 417, 423, 430, 448, 476, 532, 540, 556, 576, 623, 645, 1213, 1255, 2041, 2077, 22

Query Match 83.2%; Score 328; DB 1; Length 3011;
Best Local Similarity 76.9%; Pred. No. 1.2e-26;
Matches 60; Conservative 11; Mismatches 7; Indels 0; Gaps 0;
Qy 2 VVIVGRITLGRPAVDPREVLYQFDEMEBCASHLPYIEQGMQLAEQFKKALGLLQTA 61
Db 1680 VVIVGRVLSGKPAIDPREVLYQFDEMEBCASHLPYIEQGMQLAEQFKKALGLLQTA 1739
Qy 62 TKQAEAAAPVVEKWRAL 79
Db 1740 SRQAEVIAPVQTNNQKL 1757
RESULT 9
GNWVCH
polyprotein - hepatitis C virus (strain H)
N;Contents: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A36814; A41546
R;Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
submitted to GenBank, July 1992
A;Description: Genomic structure of the human prototype strain H of hepatitis C virus: C
A;Reference number: A36814
A;Accession: A36814
A;Molecule type: genomic RNA
A;Residues: 1-3011 <INC>
A;Cross-references: UNIPROT:P27958; GB:M67463; PIDN:AAA45534.1; PID:g329738
R;Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A;Title: Genomic structure of the human prototype strain H of hepatitis C virus: compari
A;Reference number: A41546; MUID:92052256; PMID:1658800
A;Contents: annotation

A;Note: neither amino acid nor nucleotide sequence is given
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
F;1-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;730-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;1007-1615/Product: nonstructural protein NS2 #status predicted <NS2>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F;196, 209, 234, 305, 325, 417, 423, 430, 448, 476, 532, 540, 556, 576, 623, 645, 1213, 1255, 2041, 2240, 23

Query Match 82.2%; Score 324; DB 1; Length 3011;
Best Local Similarity 76.9%; Pred. No. 3.1e-26;
Matches 60; Conservative 9; Mismatches 9; Indels 0; Gaps 0;
Qy 2 VVIVGRITLGRPAVDPREVLYQFDEMEBCASHLPYIEQGMQLAEQFKKALGLLQTA 61
Db 1680 VVIVGRVLSGKPAIDPREVLYQFDEMEBCASHLPYIEQGMQLAEQFKKALGLLQTA 1739
Qy 62 TKQAEAAAPVVEKWRAL 79
Db 1740 SRQAEVIAPVQTNNQKL 1757
RESULT 10
JQ1366
polyprotein - hepatitis C virus (French isolate) (fragments)
C;Species: hepatitis C virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: JQ1366
R;Kremsdorff, D.; Porchon, C.; Kim, J.P.; Reyes, G.R.; Brechot, C.
J. Gen. Virol. 72, 2557-2561, 1991
A;Title: Partial nucleotide sequence analysis of a French hepatitis C virus: implication
A;Reference number: JQ1366; MUID:92013977; PMID:1655961
A;Accession: JQ1366
A;Molecule type: genomic RNA
A;Residues: 1-716 <KRE>
A;Cross-references: UNIPROT:Q9PKZ2
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: glycoprotein; polyprotein
F;84, 90, 97, 115, 143, 199, 223, 243, 290, 312/Binding site: carbohydrate (Asn) (covalent) #stat

Query Match 81.2%; Score 320; DB 2; Length 716;
Best Local Similarity 75.6%; Pred. No. 1.8e-26;
Matches 59; Conservative 10; Mismatches 9; Indels 0; Gaps 0;
Qy 2 VVIVGRITLGRPAVDPREVLYQFDEMEBCASHLPYIEQGMQLAEQFKKALGLLQTA 61
Db 629 VVIVGRVLSGKPAIDPREVLYQFDEMEBCASHLPYIEQGMQLAEQFKKALGLLQTR 688
Qy 62 TKQAEAAAPVVEKWRAL 79
Db 689 SRQAEVIAPVQTNNQRL 706
RESULT 11
JC5620
genome polyprotein - hepatitis C virus (isolate EUH1480)
N;Contents: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: JC5620
R;Chamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predominant
A;Reference number: JC5620; MUID:97366593; PMID:9223423

A:Accession: JCS620

A:Molecule type: mRNA

A:Residues: 1-3014 <CHA>

A:Cross-references: UNIPROT:Q39928; GB:Y13184

A:Note: The translation of the nucleotide sequence is not complete in this paper

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin

F:2-115/Product: capsid protein C #status predicted <CPC>

F:116-191/Product: envelope protein M #status predicted <EPM>

F:192-389/Product: major envelope protein E #status predicted <ME>

F:384-408/Product: hypervariable #status predicted

F:390-730/Product: nonstructural protein NS1 #status predicted <NS1>

F:731-1007/Product: nonstructural protein NS2 #status predicted <NS2>

F:1008-1616/Product: hepatitis C virus genome polyprotein

F:1231-1238/Region: nucleotide-binding motif A (P-loop)

F:1313-1318/Region: nucleotide-binding motif B

F:1317-1320/Region: DEXH motif

F:1617-1863/Product: nonstructural protein NS4a #status predicted <NS4>

F:1864-2014/Product: nonstructural protein NS4b #status predicted <NS4b>

F:2015-3014/Product: nonstructural protein NS5 #status predicted <NS5>

F:2210-2249/Region: interferon sensitivity determining #status predicted

Query Match 64.5%; Score 254; DB 1; Length 3014;

Best Local Similarity 61.8%; Pred. No. 1.1e-18;

Matches 47; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

QY 1 SVVIGRIILSGRPVIPPDEVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLLOT 60

DB 1680 SVAIVGRIILSGRPVIPPDEVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLLOT 1739

QY 61 ATKQAEAAAAPVVEKRWAL 76

DB 1740 AQCKAETLKPATSMW 1755

RESULT 12

genome polyprotein NS4a epitope containing region (isolate HD10-1) - hepatitis C virus

C:Species: hepatitis C virus

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004

A:Accession: PC1307

R:Stuyver, L.; Van Arnhem, W.; Wyseur, A.; Deleys, R.; Maertens, G.

Biochem. Biophys. Res. Commun. 192, 635-641, 1993

A:Title: Analysis of the putative E1 envelope and NS4a epitope regions of HCV type 3.

A:Reference number: PC1300; MUID:93249436; PMID:7683463

A:Accession: PC1307

A:Molecule type: mRNA

A:Residues: 1-142 <STU>

A:Cross-references: UNIPROT:Q68870; DBJ:D14602

A:Experimental source: blood

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: polyprotein

Query Match 62.7%; Score 247; DB 2; Length 142;

Best Local Similarity 59.0%; Pred. No. 2.2e-19;

Matches 46; Conservative 12; Mismatches 20; Indels 0; Gaps 0;

QY 2 VVIVGRIILSGRPVIPPDEVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLLOT 61

DB 58 VVIVGRIILSGRPVIPPDEVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLLOT 117

QY 62 TKQAEAAAAPVVEKRWAL 79

DB 118 TQQAIVIEPIVTNWQKL 135

RESULT 13

genome polyprotein NS4a epitope containing region (isolate BR36-20) - hepatitis C virus

C:Species: hepatitis C virus

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004

A:Accession: PC1306

R:Stuyver, L.; Van Arnhem, W.; Wyseur, A.; Deleys, R.; Maertens, G.

Biochem. Biophys. Res. Commun. 192, 635-641, 1993

A:Title: Analysis of the putative E1 envelope and NS4a epitope regions of HCV type 3.

A:Reference number: PC1300; MUID:93249436; PMID:7683463

A:Accession: PC1306

A:Molecule type: mRNA

A:Residues: 1-209 <STU>

A:Cross-references: UNIPROT:Q81594; DBJ:D14600; NID:G303584; PIDN:BA003449.1; PID:G3035

A:Experimental source: blood

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: nonstructural protein; polyprotein

Query Match 62.7%; Score 247; DB 2; Length 209;

Best Local Similarity 57.7%; Pred. No. 3.4e-19;

Matches 45; Conservative 13; Mismatches 20; Indels 0; Gaps 0;

QY 2 VVIVGRIILSGRPVIPPDEVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLLOT 61

DB 125 VVIVGRIILSGRPVIPPDEVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLLOT 184

QY 62 TKQAEAAAAPVVEKRWAL 79

DB 185 TQQAIVIEPIVTNWQKL 202

RESULT 14

PC2219

polyprotein - hepatitis C virus (type 5a) (fragments)

N:Contains: core protein; E1 (carboxyl end); E2/NS1 (amino end); NS3 protein; NS4A prot

C:Species: hepatitis C virus

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004

A:Accession: PC2219

R:Stuyver, L.; Arnhem, W.V.; Wyseur, A.; Maertens, G.

Biochem. Biophys. Res. Commun. 202, 1308-1314, 1994

A:Title: Cloning and phylogenetic analysis of the core, E2, and NS3/NS4 regions of the

A:Reference number: PC2219; MUID:94338342; PMID:7520237

A:Accession: PC2219

A:Molecule type: mRNA

A:Residues: 1-876 <STU>

A:Cross-references: UNIPROT:Q81242; GB:L29577; GB:L29578; GB:L29579

A:Experimental source: serum

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: glycoprotein

F:1-191/Product: core #status predicted <COE>

F:68-78/Region: variable

F:192-247/Product: E1 (carboxyl end) #status predicted <ERE>

F:248-411/Product: E2/NS1 (amino end) #status predicted <ENR>

F:248-338/Region: E2

F:339-411/Region: NS1 (amino end)

F:412-783/Product: NS3 #status predicted <NSR>

F:784-837/Product: NS4A #status predicted <NSA>

F:838-876/Product: NS4B #status predicted <NSB>

F:281,287,294,312,340/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 58.1%; Score 229; DB 2; Length 876;

Best Local Similarity 60.0%; Pred. No. 1.4e-16;

Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 SVVIGRIILSGRPVIPPDEVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLLOT 60

DB 805 SVAIVGRIILSGRPVIPPDEVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLLOT 864

QY 61 ATKQAEAAAAP 70

DB 865 TQQAIVIEPIVTNWQKL 874

RESULT 15

GNWVJ8

genome polyprotein - hepatitis C virus (strain HC-J8)

N:Contains: capsid protein C; envelope protein M; hepatitis C virus (nonstruct

protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

C:Species: hepatitis C virus

Sat Aug 13 10:39:23 2005

C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A40250; PQ0397; PQ0559
R;Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fukuda, S.;
Virology 188, 331-341, 1992
A;Title: Full-length sequence of a hepatitis C virus genome having poor homology to repo
A;Reference number: A40250; MUID:92230232; PMID:1314459
A;Accession: A40250
A;Molecule type: genomic RNA
A;Residues: 1-3033 <OKA>
A;Cross-references: UNIPROT:P26661; GB:D10988; GB:D01221; NID:g221608; PIDN:BAA01761.1;
R;Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.I.
J. Gen. Virol. 73, 1131-1141, 1992
A;Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e
A;Reference number: PQ0393; MUID:92268871; PMID:1316939
A;Accession: PQ0397
A;Molecule type: genomic RNA
A;Residues: 2678-2754 <CHA>
A;Cross-references: DBJ:D10134
A;Experimental source: isolate B-b12
R;kato, N.; Ootsuyama, Y.; Onkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohnc
Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A;Title: Distribution of plural HCV types in Japan.
A;Reference number: PQ0554; MUID:92068204; PMID:1720309
A;Accession: PQ0559
A;Molecule type: mRNA
A;Residues: 2678-2729 <KAT>
A;Cross-references: GB:D10562; GB:D90518; NID:g221523; PIDN:BAA01418.1; PID:g221524
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
F;1-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-733/Product: nonstructural protein NS1 #status predicted <NS1>
F;734-1010/Product: nonstructural protein NS2 #status predicted <NS2>
F;1011-1619/Product: hepatitis C virus #status predicted <NS3>
F;1234-1241/Region: nucleotide-binding motif A (P-loop)
F;1316-1321/Region: nucleotide-binding motif B
F;1320-1323/Region: DEXH motif
F;1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
F;2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>
F;196, 209, 233, 299, 305, 417, 423, 430, 448, 477, 534, 542, 558, 578, 627, 649, 1091, 1217, 1259, 2038, 23

Query Match 52.8%; Score 208; DB 1; Length 3033;
Best Local Similarity 50.0%; Pred. No. 9.5e-14;
Matches 39; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

QY 2 VVIVGRITLGRPAVIPDREVLVYQEFDEMEECASHLPVIEQGMOLAEQFKOKALGLLOTA 61
DB 1684 ISIIIGRLHNDRVVVPDKIYLIEAFDEMEECASKAALIEGQRMELKSKIQGLLOQA 1743

QY 62 TKQAEAAAPVVESKWRAL 79
DB 1744 TRQAQDIQPAIQSSWPKL 1761

Search completed: August 12, 2005, 14:22:49
Job time : 17.7224 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:03:18 ; Search time 72.9828 Seconds
(without alignments)
554.298 Million cell updates/sec

Title: US-09-758-308-3

Perfect score: 394

Sequence: 1 SWIVIGRIILSRPAVIPDR.....TATKQAEAAAPVVEKMRAL 79

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	394	100.0	3010	1	POLG HCVJA
2	394	100.0	3010	2	P26662 h genome po
3	394	100.0	3010	2	Q93077 hepatitis c
4	394	100.0	3010	2	Q68788 hepatitis c
5	394	100.0	3010	2	Q81757 hepatitis c
6	394	100.0	3013	2	Q636P5
7	393	99.7	138	2	Q9QIX9
8	393	99.7	3013	2	Q68217
9	392	99.5	3010	2	Q9QIY0
10	391	99.2	3010	2	Q9DTE5
11	391	99.2	3010	2	Q9QIX5
12	390	99.0	138	2	Q68218
13	390	99.0	659	2	Q68K36
14	390	99.0	659	2	Q68K57
15	390	99.0	659	2	Q68K69
16	390	99.0	3010	2	Q913V3
17	390	99.0	3010	2	Q9DTE6
18	390	99.0	3010	2	Q9DTE7
19	390	99.0	3010	2	Q9J3G8
20	390	99.0	3010	2	Q9J3H1
21	390	99.0	3010	2	Q9J3H3
22	390	99.0	3010	2	Q9J3H8
23	390	99.0	3010	2	Q9QIY3
24	390	99.0	3010	2	Q9QIY4
25	390	99.0	3010	2	Q9QIY9
26	390	99.0	3010	2	Q9QIZ0
27	390	99.0	3013	2	Q9J3H4
28	389	98.7	138	2	Q68205
29	389	98.7	138	2	Q68244
30	389	98.7	3010	1	POLG HCVBK
31	389	98.7	3010	2	Q9DTE2

32	389	98.7	3010	2	Q9DTE8	Q9dte8 hepatitis c
33	389	98.7	3010	2	Q9QPE1	Q9qp61 hepatitis c
34	389	98.7	3015	2	Q9WPH5	Q9wph5 hepatitis c
35	388	98.5	138	2	Q68216	Q68216 hepatitis c
36	388	98.5	3010	2	Q68533	Q68533 hepatitis c
37	387	98.2	138	2	Q68215	Q68215 hepatitis c
38	387	98.2	138	2	Q68227	Q68227 hepatitis c
39	387	98.2	138	2	Q68228	Q68228 hepatitis c
40	387	98.2	659	2	Q68K50	Q68K50 hepatitis c
41	387	98.2	3010	2	Q91AU0	Q91au0 hepatitis c
42	387	98.2	3010	2	Q807P3	Q807p3 hepatitis c
43	387	98.2	3010	2	Q9DTE0	Q9dtf0 hepatitis c
44	387	98.2	3010	2	Q9J3G5	Q9j3g5 hepatitis c
45	387	98.2	3010	2	Q9J3H0	Q9j3h0 hepatitis c

ALIGNMENTS

RESULT 1

POLG_HCVJA STANDARD; PRT; 3010 AA.
AC P26662;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Genome polyprotein [Contains: Capsid protein C (Core protein) (P22); Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21) (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
OS Hepatitis C virus (isolate Japanese) (HCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
OC NCBI_TaxID=11116;
RP SEQUENCE FROM N.A.
RX MEDLINE=91088550; PubMed=2175903;
RA Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S., Sugimura T., Shimotohno K.;
RT "Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
RN [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=91192160; PubMed=1849488; DOI=10.1016/0014-5793(91)80322-T;
RA Kato N., Hijikata M., Nakagawa M., Ootsuyama Y., Muraishi K., Ohkoshi S., Shimotohno K.;
RT "Molecular structure of the Japanese hepatitis C viral genome.";
RL FEBS Lett. 280:325-328(1991).
CC -!- FUNCTION: The small proteins NS2A, NS2B, NS4A and NS4B are hydrophobic, suggesting a possible membrane-related function. NS3 and NS5 may play a role in the viral RNA replication.
CC and NS5 may play a role in the viral RNA replication.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
CC -!- CATALYTIC ACTIVITY: Nucleoside triphosphate = N diphosphate + {RNA} (N).
CC -!- SUBUNIT: The virion of this virus is a nucleocapsid covered by a lipoprotein envelope. The envelope consists of two proteins: protein M and glycoprotein E. The nucleocapsid is a complex of protein C and mRNA.
CC -!- SIMILARITY: Contains 1 peptidase S29 domain.
CC -!- SIMILARITY: Contains 1 peptidase U39 domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

```
CC or send an email to license@iab-sib.ch).
CC -----
DR EMBL; D90208; BAA14233.1; -.
DR PIR; A39253; GNWVCJ.
DR HSSP; P26663; 1JXP.
DR MEROPS; S29.001; -.
DR MEROPS; U39.001; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral_RdRP; 1.
DR SMART; SMO0487; DEXDC; 1.
KW ATP-binding; Coat protein; Core protein; Core protein; Envelope protein;
KW Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
KW RNA-directed RNA polymerase; Serine protease; Transferase;
KW Transmembrane.
FT INIT_MET 1 1 Removed from capsid protein C by the
FT CHAIN 1 115 cellular aminopeptidase.
FT CHAIN 116 191 Capsid protein C (Potential).
FT CHAIN 192 283 Matrix protein (Potential).
FT CHAIN 384 729 Major envelope protein E (Potential).
FT CHAIN 730 1006 Nonstructural protein NS1 (Potential).
FT CHAIN 1007 1615 Nonstructural protein NS2 (Potential).
FT CHAIN 1616 1862 Protease/helicase NS3 (Potential).
FT CHAIN 1863 2013 Nonstructural protein NS4A (Potential).
FT CHAIN 2014 3010 Nonstructural protein NS4B (Potential).
FT TRANSMEM 347 369 RNA-directed RNA polymerase (Potential).
FT ACT_SITE 1083 1083 Potential.
FT ACT_SITE 1107 1107 Charge relay system (By similarity).
FT ACT_SITE 1165 1165 Charge relay system (By similarity).
FT NP_BIND 1230 1237 ATP (Potential).
FT SITE 1316 1319 DECH box.
FT CARBOHYD 196 196 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 209 209 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 234 234 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 250 250 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 305 305 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 417 417 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 423 423 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 430 430 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 448 448 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 532 532 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 556 556 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 576 576 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 623 623 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 645 645 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 2041 2041 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 2077 2077 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 2240 2240 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 2788 2788 N-linked (GlcNAc...) (Potential).

SQ SEQUENCE 3010 AA; 327017 MW; AA993794F46DB185 CRC64;
Query Match 100.0%; Score 394; DB 1; Length 3010;
Best Local Similarity 100.0%; Pred. No. 4.le-32;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SWIVGRILISGRPAVIPDREVLVQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLOT 60
Db 1679 SWIVGRILISGRPAVIPDREVLVQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLOT 1738
QY 61 ATKQAEAAAPVVESKWRAL 79
Db 1739 ATKQAEAAAPVVESKWRAL 1757

RESULT 2
O93077 PRELIMINARY; PRT; 3010 AA.
AC O93077;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98122498; PubMed=9462666; DOI=10.1002/hep.510270242;
RA Aizaki H., Aoki Y., Harada T., Ishii K., Suzuki T., Nagamori S.,
RA Toda G., Matsuura Y., Miyamura T.,
RT "Full-length complementary DNA of hepatitis C virus genome from an
RL Hepatology 27:621-627(1998).
DR EMBL; D89815; BAA25076.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CWX.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005138; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transmembrane; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
```

```

DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02307; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; Viral RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327352 MW; 888BBA102A733390 CRC64;

Query Match 100.0%; Score 394; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 4.1e-32;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVVIVGRILSGRPAVIPDREVLVYQEFDEMEECSAHLPIYIEQGMQLAEQFKQKALGLLOT 60
Db 1679 SVVIVGRILSGRPAVIPDREVLVYQEFDEMEECSAHLPIYIEQGMQLAEQFKQKALGLLOT 1738

QY 61 ATKQAEAAAPVVESKWRAL 79
Db 1739 ATKQAEAAAPVVESKWRAL 1757

RESULT 3
Q68788 PRELIMINARY; PRT; 3010 AA.
AC Q68788;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE HCV polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96362158; PubMed=8720135;
RA Seki M., Honda Y.;
RT "phosphorocholate antisense oligodeoxynucleotides capable of
inhibiting hepatitis C virus gene expression: in vitro translation
assay."
RL J. Biochem. 118:1199-1204(1995).
DR EMBL; D45172; BAA08120.1; -
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q81755; IDXP.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:008236; F:serine-type peptidase activity; IEA.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:000508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:vital genome replication; IEA.
DR GO; GO:0019087; P:vital transformation; IEA.
DR InterPro; IPR00345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV capsid.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV_NSI.

```

DR GO: 0005198; F: structural molecule activity; IEA.
 DR GO: 0006508; P: proteolysis and peptidolysis; IEA.
 DR GO: 0006350; P: transcription; IEA.
 DR GO: 0019079; P: viral genome replication; IEA.
 DR GO: 0019087; P: viral transformation; IEA.
 DR InterPro: IPR000345; CytC_heme_BS.
 DR InterPro: IPR001410; DEAD/DEAH_N.
 DR InterPro: IPR011545; DEAD/DEAH_N.
 DR InterPro: IPR002522; HCV capsid.
 DR InterPro: IPR002521; HCV core.
 DR InterPro: IPR002519; HCV env.
 DR InterPro: IPR002531; HCV NS1.
 DR InterPro: IPR000745; HCV NS4a.
 DR InterPro: IPR001490; HCV NS4b.
 DR InterPro: IPR002868; HCV NS5a.
 DR InterPro: IPR002166; HCV RdRP.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR004109; Peptidase_S29.
 DR InterPro: IPR009003; Pept_Ser_Cys.
 DR InterPro: IPR002518; Pept_U39_HCV_NS2.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR007094; RNA_pol_PSVir.
 DR Pfam: PF01543; HCV capsid; 1.
 DR Pfam: PF01542; HCV core; 1.
 DR Pfam: PF01539; HCV env; 1.
 DR Pfam: PF01560; HCV NS1; 1.
 DR Pfam: PF01538; HCV NS2; 1.
 DR Pfam: PF02907; HCV NS3; 1.
 DR Pfam: PF01006; HCV NS4a; 1.
 DR Pfam: PF01001; HCV NS4b; 1.
 DR Pfam: PF01506; HCV NS5a; 1.
 DR Pfam: PF00271; Helicase_C; 1.
 DR Pfam: PF00998; Viral_RdRP; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 SQ SEQUENCE 3010 AA; 326820 MW; 49C7CA6E22468523 CRC64;

Query Match 100.0%; Score 394; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 4.1e-32;
 Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVWIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLARQFKQKALGLLOT 60
 |||||
 Db 1679 SVWIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLARQFKQKALGLLOT 1738

QY 61 ATKQAEAAAPVVEKWRAL 79
 |||||
 Db 1739 ATKQAEAAAPVVEKWRAL 1757

RESULT 5
 Q6JG65 PRELIMINARY; PRT; 3013 AA.
 AC Q6JG65;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Polyprotein (fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]_TaxID=11103;
 RP SEQUENCE FROM N.A.
 RA rao X., Guo J., Zheng C.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Li X., Zheng C.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY587016; AAT40682.1; --

DR HSPSP; P26663; IUXP.
 DR GO: 0016021; C: integral to membrane; IEA.
 DR GO: 0019028; C: viral capsid; IEA.
 DR GO: 0019031; C: viral envelope; IEA.
 DR GO: 0005524; F: ATP binding; IEA.
 DR GO: 0008026; F: ATP-dependent helicase activity; IEA.
 DR GO: 0003723; F: RNA binding; IEA.
 DR GO: 0003968; F: RNA-directed RNA polymerase activity; IEA.
 DR GO: 0008236; F: serine-type peptidase activity; IEA.
 DR GO: 0005198; F: structural molecule activity; IEA.
 DR GO: 0006508; P: proteolysis and peptidolysis; IEA.
 DR GO: 0006350; P: transcription; IEA.
 DR GO: 0019079; P: viral genome replication; IEA.
 DR GO: 0019087; P: viral transformation; IEA.
 DR InterPro: IPR000345; CytC_heme_BS.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR011545; DEAD/DEAH_N.
 DR InterPro: IPR002522; HCV capsid.
 DR InterPro: IPR002521; HCV core.
 DR InterPro: IPR002519; HCV env.
 DR InterPro: IPR002531; HCV NS1.
 DR InterPro: IPR000745; HCV NS4a.
 DR InterPro: IPR001490; HCV NS4b.
 DR InterPro: IPR002868; HCV NS5a.
 DR InterPro: IPR002166; HCV RdRP.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR009003; Pept_Ser_Cys.
 DR InterPro: IPR002518; Pept_U39_HCV_NS2.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR007094; RNA_pol_PSVir.
 DR Pfam: PF01543; HCV capsid; 1.
 DR Pfam: PF01542; HCV core; 1.
 DR Pfam: PF01539; HCV env; 1.
 DR Pfam: PF01560; HCV NS1; 1.
 DR Pfam: PF01538; HCV NS2; 1.
 DR Pfam: PF01006; HCV NS4a; 1.
 DR Pfam: PF01001; HCV NS4b; 1.
 DR Pfam: PF01506; HCV NS5a; 1.
 DR Pfam: PF00271; Helicase_C; 1.
 DR Pfam: PF00998; Viral_RdRP; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 FT NON_TER 3013 3013
 SQ SEQUENCE 3013 AA; 327180 MW; C71FA2B7C5257F3D CRC64;

Query Match 100.0%; Score 394; DB 2; Length 3013;
 Best Local Similarity 100.0%; Pred. No. 4.1e-32;
 Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVWIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLARQFKQKALGLLOT 60
 |||||
 Db 1679 SVWIVGRILSGRPAVDPREVLYQEFDEMEECASHLPYIEQGMOLARQFKQKALGLLOT 1738

QY 61 ATKQAEAAAPVVEKWRAL 79
 |||||
 Db 1739 ATKQAEAAAPVVEKWRAL 1757

RESULT 6
 Q6QIX9 PRELIMINARY; PRT; 3013 AA.
 AC Q6QIX9;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]_TaxID=11103;

RP SEQUENCE FROM N.A.
RC STRAIN=MD6-2;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.I., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity";
RL Virology 263:244-253(1999).
DR EMBL; AF165056; AAD56191.1; --
DR PIR; A61196; A61196.
DR PIR; PQ0251; PQ0251.
DR PIR; PQ0252; PQ0252.
DR PIR; PQ0253; PQ0253.
DR PIR; PQ0254; PQ0254.
DR PIR; PQ0255; PQ0255.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; ICWX.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; F: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; F: transcription; IEA.
DR GO; GO:0019079; F: viral genome replication; IEA.
DR GO; GO:0019087; F: viral translocation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD_heme_BS.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; transmembrane.
SQ SEQUENCE 3013 AA; 326920 MW; 98D0BDE208A9B90E CRC64;

Query Match 100.0%; Score 394; DB 2; Length 3013;
Best Local Similarity 100.0%; Pred. No. 4, 1e-32;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVVIVGRIILSGRPVAPIDPREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLOT 60
Db 1682 SVVIVGRIILSGRPVAPIDPREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLOT 1741

QY 61 ATKQAEAAAPVVESKWRAL 79
Db 1742 ATKQAEAAAPVVESKWRAL 1760

RESULT 7
Q68217 ID Q68217 PRELIMINARY; PRT; 138 AA.
AC Q68217;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (fragment).
OS Name=ns4;
GN Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14253; AAC53942.1; --
DR HSP; P26663; 1CUL.
DR GO; GO:0019012; C: virion; IEA.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
KW Nonstructural protein.
FT NON_TER 1
FT NON_TER 138 138
SQ SEQUENCE 138 AA; 15130 MW; 2AF1E92DDC7B741D CRC64;

Query Match 99.7%; Score 393; DB 2; Length 138;
Best Local Similarity 98.7%; Pred. No. 2e-33;
Matches 78; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVVIVGRIILSGRPVAPIDPREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLOT 60
Db 53 SVVIVGRIILSGRPVAPIDPREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLOT 112

QY 61 ATKQAEAAAPVVESKWRAL 79
Db 113 ATKQAEAAAPVVESKWRAL 131

RESULT 8
Q9QIY0 ID Q9QIY0 PRELIMINARY; PRT; 3013 AA.
AC Q9QIY0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MD6-1;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.I., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity";
RL Virology 263:244-253(1999).
DR EMBL; AF165055; AAD56190.1; --

```

DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0251; PQ0251.
DR PIR; PQ0254; PQ0254.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CW.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0008236; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral_RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3013 AA; 326887 MW; 762E2D4B6B607B8C CRC64;

Query Match 99.7%; Score 393; DB 2; Length 3013;
Best Local Similarity 98.7%; Pred. No. 5.3e-32;
Matches 78; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SWVTVGRILSRPAVDPREVLYQFDEMEECASHLPYIEQGMOLAEQFKQKALGLLQ 60
Db 1682 SWVTVGRILSRPAIIPREVLYQFDEMEECASHLPYIEQGMOLAEQFKQKALGLLQ 1741

Qy 61 ATKQAEAAAPVVEKWRAL 79
Db 1742 ATKQAEAAAPVVEKWRAL 1760

RESULT 9
Q9DTE5 PRELIMINARY; PRT; 3010 AA.
AC Q9DTE5;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

```

```

DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA Mishiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RT with hepatocellular carcinoma: the 'progression score' revisited.";
RL Hepatol. Res. 20:161-171(2001).
EMBL; AB049092; BAB1805.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CW.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0008236; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral_RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 2.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 326987 MW; 573C0F3C55B3F3F4 CRC64;

Query Match 99.5%; Score 392; DB 2; Length 3010;
Best Local Similarity 97.5%; Pred. No. 6.7e-32;
Matches 77; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SWVTVGRILSRPAVDPREVLYQFDEMEECASHLPYIEQGMOLAEQFKQKALGLLQ 60

```


Db 1679 SVIVGRIVLGRPAVVPDREVLVYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLOT 1738

61 ATKOAEAAAPVESKWRAL 79

db 1739 ATKOAEAAA PVVESKWRAL 1757

RESULT 10

ID	Q9DTE4	PRELIMINARY	PRT	3010 AA.
AC	Q9DTE4			
DT	01-MAR-2001	(TrEMBLrel. 16, Created)		
DT	01-MAR-2001	(TrEMBLrel. 16, Last sequence update)		
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)		
DE		Polyprotein.		
OS		Hepatitis C virus.		
OC		Viruses; SERNA positive-strand viruses, no DNA stage; Flaviviridae;		
CC		Hepacivirus.		
OX		NCBI_TaxID=11103;		
RN	[1]	SEQUENCE FROM N.A.		
RP		TISSUE=Serum;		
RC		Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,		
RA		Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,		
RA		Mishiro S.;		
RT		"Hepatitis C virus (HCV) genotype 1b sequences from fifteen patient		
RT		with hepatocellular carcinoma: the 'progression score' revisited.";		
RL		Hepatol. Res. 20:161-171(2001).		
DR	ENBL	AB049093; BAB18806.1; -.		
DR	PIR	A61196; A61196.		
DR	PIR	PQ0246; PQ0246.		
DR	PIR	PQ0804; PQ0804.		
DR	PIR	PS0329; PS0329.		
DR	HSSP	Q8Y5I; ICWX.		
DR	GO	GO:0016021; C:integral to membrane; IEA.		
DR	GO	GO:0019028; C:viral capsid; IEA.		
DR	GO	GO:0019031; C:viral envelope; IEA.		
DR	GO	GO:0005524; F:ATP binding; IEA.		
DR	GO	GO:0008026; F:ATP-dependent helicase activity; IEA.		
DR	GO	GO:0003723; F:RNA binding; IEA.		
DR	GO	GO:0003968; F:RNA-directed RNA polymerase activity; IEA.		
DR	GO	GO:0008236; F:serine-type peptidase activity; IEA.		
DR	GO	GO:0005198; F:structural molecule activity; IEA.		
DR	GO	GO:0006508; P:proteolysis and peptidolysis; IEA.		
DR	GO	GO:0006350; P:transcription; IEA.		
DR	GO	GO:0019079; P:viral genome replication; IEA.		
DR	GO	GO:0019087; P:viral transformation; IEA.		
DR	InterPro	IPR000345; CytC_heme_BS.		
DR	InterPro	IPR001410; DEAD.		
DR	InterPro	IPR011545; DEAD/DEAH_N.		
DR	InterPro	IPR002522; HCV_capsid.		
DR	InterPro	IPR002521; HCV_core.		
DR	InterPro	IPR002519; HCV_env.		
DR	InterPro	IPR002531; HCV_NS1.		
DR	InterPro	IPR000745; HCV_NS4a.		
DR	InterPro	IPR001490; HCV_NS4b.		
DR	InterPro	IPR002868; HCV_NS5a.		
DR	InterPro	IPR002166; HCV_RDRP.		
DR	InterPro	IPR001650; Helicase_C.		
DR	InterPro	IPR004109; Peptidase_S29.		
DR	InterPro	IPR009003; Pept_Ser_Cys.		
DR	InterPro	IPR002518; Pept_u39_HCV_NS2.		
DR	InterPro	IPR007095; RNA_pol_DS_P5.		
DR	InterPro	IPR007094; RNA_pol_PSVir.		
DR	Pfam	PF01543; HCV_capsid; 1.		
DR	Pfam	PF01542; HCV_core; 1.		
DR	Pfam	PF01539; HCV_env; 1.		
DR	Pfam	PF01560; HCV_NS1; 1.		
DR	Pfam	PF01538; HCV_NS2; 1.		
DR	Pfam	PF02307; HCV_NS3; 1.		
DR	Pfam	PF01006; HCV_NS4a; 1.		
DR	Pfam	PF01001; HCV_NS4b; 1.		
DR	Pfam	PF01506; HCV_NS5a; 1.		

DR pfam; PF00271; Helicase C; 1.
DR pfam; PF00598; Viral_RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
DR KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein
KW Polyprotein; Transmembrane.
DR SEQUENCE 3010 AA; 327325 MW; 3D8CF249BD1151C CRC64;

Query Match 99.2% Score 391 DB 2 Length 3010;

Best Local Similarity	98.7%;	Pred. No. 8.5e-32;
Matches	78:	Conservative
Matches	1:	Mismatches
Indels	0:	Indels
Gaps	0:	Gaps

Qy	1	SVWVGR	II	SGRP	PAV	TPD	REVL	YQ	FEDE	ME	EC	ASH	LPY	IEQ	GM	LA	EQFK	KAL	GL	LOT	60
Nb	1679	SVWVGR	II	SGK	PAV	TPD	REVL	YQ	FEDE	ME	EC	ASH	LPY	IEQ	GM	LA	EQFK	KAL	GL	LOT	1738

61 ATKOAEAAAPVVEKWRAL, 79

1739 ATKQAEAAAPVVESKRAL 1757

RESIT.T 11

Q9QIX5	PRELIMINARY; PRT; 3010 AA.
ID	Q9QIX5
AC	Q9QIX5; (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE	Polyprotein.
OS	Hepatitis C virus.
OC	Hepacivirus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC	Hepacivirus.
OX	NCBI_TaxID=11103;
OX	[1]_TaxID=11103;
RN	SEQUENCE FROM N.A.
RP	STRAIN=MD8-2;
RC	MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RC	MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA	Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA	Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT	"Time-related changes in full-length hepatitis C virus and hepatitis
RT	activity.";
RT	activity.";
RL	Virology 283:244-253(1999).
DR	EMBL; AF165060; AAD56195.1; -.
DR	PIR; A61196; A61196.
DR	PIR; PQ0246; PQ0246.
DR	PIR; PQ0254; PQ0254.
DR	PIR; PQ0804; PQ0804.
DR	PIR; PS0329; PS0329.
DR	HSP; Q8JYS1; ICWX.
DR	MEROPS; S29.002; -.
DR	GO; GO:0016021; C:integral to membrane; IEA.
DR	GO; GO:0013028; C:viral capsid; IEA.
DR	GO; GO:0013031; C:viral envelope; IEA.
DR	GO; GO:0005524; F:ATP binding; IEA.
DR	GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR	GO; GO:0003723; F:RNA binding; IEA.
DR	GO; GO:0003969; F:RNA-directed RNA polymerase activity; IEA.
DR	GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR	GO; GO:0005198; F:structural molecule activity; IEA.
DR	GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR	GO; GO:0006350; F:transcription; IEA.
DR	GO; GO:0013079; F:viral genome replication; IEA.
DR	GO; GO:0013087; F:viral transformation; IEA.
DR	InterPro; IPR000345; CytC_heme_BS.
DR	InterPro; IPR001410; DEAD.
DR	InterPro; IPR011545; DEAD/DEAH N.
DR	InterPro; IPR002522; HCV capsid.
DR	InterPro; IPR002521; HCV core.
DR	InterPro; IPR002519; HCV env.
DR	InterPro; IPR002531; HCV NS1.
DR	InterPro; IPR000745; HCV NS4a.
DR	InterPro; IPR001490; HCV NS4b.
DR	InterPro; IPR002868; HCV NS5a.

```

DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral_RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327295 MW; 8B99F1BBA6A50F56 CRC64;

Query Match 99.2%; Score 391; DB 2; Length 3010;
Best Local Similarity 98.7%; Pred. No. 8.5e-32;
Matches 78; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVTVGRILSGRAVDPREVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLIQT 60
DB 1679 SVTVGRILSGRAVDPREVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLIQT 1738

QY 61 ATKQAEAAAPVVESKWRAL 79
DB 1739 ATKQAEAAAPVVESKWRAL 1757

RESULT 12
ID Q68218 PRELIMINARY; PRT; 138 AA.
AC Q68218;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheng M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14254; AAC53943.1; -.
DR HSSP; P26663; 1CUL.
DR GO; GO:0019012; C:virion; IEA.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
KW Nonstructural protein.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15189 MW; DB78E92DDC67040F CRC64;

Query Match 99.0%; Score 390; DB 2; Length 138;
Best Local Similarity 97.5%; Pred. No. 4.2e-33;
Matches 77; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SVTVGRILSGRAVDPREVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLIQT 60
DB 53 SVTVGRILSGRAVDPREVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLIQT 112

QY 61 ATKQAEAAAPVVESKWRAL 79
DB 113 ATKQAEAAAPVVESKWRAL 131

RESULT 13
ID Q68K36 PRELIMINARY; PRT; 659 AA.
AC Q68K36;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polyprotein (fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685636; AAT94277.1; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICG; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70411 MW; 2FCC6D44A67324E1 CRC64;

Query Match 99.0%; Score 390; DB 2; Length 659;
Best Local Similarity 97.5%; Pred. No. 2.2e-32;
Matches 77; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVTVGRILSGRAVDPREVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLIQT 60
DB 574 SVTVGRILSGRAVDPREVLYQFDEMEECASHLPYIEQGMQLAEQFKKALGLIQT 633

QY 61 ATKQAEAAAPVVESKWRAL 79
DB 634 ATKQAEAAAPVVESKWRAL 652

RESULT 14
ID Q68K57 PRELIMINARY; PRT; 659 AA.
AC Q68K57;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polyprotein (fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.

```

RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
 RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
 RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
 RT HCV/HIV Coinfected Subjects.";
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY685592; AAT94256.1; -;
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR000745; HCV NS4a_N.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR004109; Peptidase_S29.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR Pfam; PF02907; HCV NS3; 1.
 DR Pfam; PF01006; HCV NS4a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELICc; 1.
 KW Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 659 659
 SQ SEQUENCE 659 AA; 70411 MW; 2FCC6D4A67324E1 CRC64;

Query Match 99.0%; Score 390; DB 2; Length 659;
 Best Local Similarity 97.5%; Pred. No. 2.2e-32;
 Matches 77; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVVIVGRILSGRPAVIPDREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLOT 60
 Db 574 SVVIVGRILSGKPAIIPDREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLOT 633
 Qy 61 ATKQAEAAAAPVVESKWRAL 79
 Db 634 ATKQAEAAAAPVVESKWRAL 652

RESULT 15
 Q68K69 PRELIMINARY; PRT; 659 AA.
 AC Q68K69
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=111103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
 RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
 RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
 RT HCV/HIV Coinfected Subjects.";
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY685559; AAT94244.1; -;
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR000745; HCV NS4a.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR004109; Peptidase_S29.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR Pfam; PF02907; HCV NS3; 1.
 DR Pfam; PF01006; HCV NS4a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELICc; 1.
 KW Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 659 659
 SQ SEQUENCE 659 AA; 70593 MW; CFF76C7E0242545 CRC64;
 Query Match 99.0%; Score 390; DB 2; Length 659;
 Best Local Similarity 97.5%; Pred. No. 2.2e-32;

Matches 77; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 SVVIVGRILSGRPAVIPDREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLOT 60
 Db 574 SVVIVGRILSGKPAIIPDREVLYQEFDEMEECASHLPYIEQGMQLAEQFKQKALGLLOT 633
 Qy 61 ATKQAEAAAAPVVESKWRAL 79
 Db 634 ATKQAEAAAAPVVESKWRAL 652

Search completed: August 12, 2005, 14:21:20
 Job time : 73.9828 secs

This Page Blank (uspto)

GemCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:00:18 ; Search time 34.6216 Seconds
(without alignments)
368.645 Million cell updates/sec

Title: US-09-758-308-4
Perfect score: 167
Sequence: 1 MRLIAFASRGHVSPTHYVPESDAARVTOIL 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 385760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	167	100.0	33	AAB31698	Aab31698 Antigenic
2	167	100.0	196	AAR25865	Aar25865 HCV poly
3	167	100.0	196	Aaw41743	Aaw41743 Hepatitis
4	167	100.0	214	ADO36214	Ado36214 Hepatitis
5	167	100.0	214	ADO79388	Ado79388 Hepatitis
6	167	100.0	219	AAR25887	Aar25887 HK12. 9/2
7	167	100.0	272	AAR29882	Aar29882 HCV NS4-N
8	167	100.0	272	AAR29883	Aar29883 HCV NS4-N
9	167	100.0	272	AAR29885	Aar29885 HCV NS4-N
10	167	100.0	272	AAR29871	Aar29871 HCV NS4-N
11	167	100.0	272	AAR29884	Aar29884 HCV NS4-N
12	167	100.0	360	AAR29877	Aar29877 HCV NS4-N
13	167	100.0	767	AAR80044	Aar80044 Hepatitis
14	167	100.0	863	AAR29881	Aar29881 HCV NS4-N
15	167	100.0	1260	AAR12599	Aar12599 Portion O
16	167	100.0	1411	AAR29533	Aar29533 HCV NS4-N
17	167	100.0	1736	AAB36932	Aab36932 Hepatitis
18	167	100.0	1985	AAO18001	Aao18001 Hepatitis
19	167	100.0	1985	AAE15729	Aae15729 Hepatitis
20	167	100.0	1985	AAE15731	Aae15731 Hepatitis
21	167	100.0	1985	AAE15720	Aae15720 Hepatitis
22	167	100.0	1985	AAE15717	Aae15717 Hepatitis
23	167	100.0	1985	AAE15727	Aae15727 Hepatitis
24	167	100.0	1985	AAE15728	Aae15728 Hepatitis
25	167	100.0	1985	AAE15722	Aae15722 Hepatitis

ALIGNMENTS

RESULT 1

AAB31698
ID AAB31698 standard; peptide; 33 AA.

AC AAB31698;

XX 30-APR-2001 (first entry)

XX Antigenic epitope of the Hepatitis C virus (HCV) NS4 protein.

XX Antigen; HCV; polyprotein; core protein; NS3 protein; NS4 protein;

KW NS4a protein; HCV infection.

XX Hepatitis C virus.

XX WO200104149-A1.

XX 18-JAN-2001.

XX 07-JUL-2000; 2000WO-US018704.

XX 09-JUL-1999; 99WO-US015578.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Fields HA, Khudyakov YE;

XX WPI; 2001-138316/14.

XX New (mosaic) polypeptides, useful as reagents in assays for the diagnosis or monitoring of HCV infections and as components of anti-HCV vaccines, comprises antigenic groups of hepatitis C virus (HCV).

XX Claim 5; Page 39; 52pp; English.

XX The present sequence represents an antigenic epitope from a hepatitis C virus (HCV) NS4 protein. The peptide comprises amino acids 1916-1948 of the HCV polyprotein. The specification describes antigenic epitopes from HCV core protein, NS3 protein, NS4 protein and NS4a protein. The antigenic peptides are useful as diagnostic reagents for detecting HCV in a biological sample. They are also useful for monitoring HCV infection in a patient sample in addition to diagnosis. Pharmaceutical compositions comprising the peptides are useful for preventing, minimizing or reducing HCV infection in patients who have been exposed to HCV or to individuals, such as health care workers or blood product recipients, who are more likely to become exposed to HCV infection. The antigenic peptides are also useful for generating antibodies which can be used to detect HCV proteins in a sample or for laboratory research purposes.

26	167	100.0	1985	5	AAE15730	Hepatitis
27	167	100.0	1985	6	ABU09574	HCV Met-N
28	167	100.0	1985	6	ABU09575	HCV Met-N
29	167	100.0	1985	8	ADJ57846	HCV repli
30	167	100.0	1985	8	ADR38450	Hepatitis
31	167	100.0	2063	7	ADD67963	Hepatitis
32	167	100.0	2201	2	AAW01680	HCV NS2-N
33	167	100.0	2201	5	ABG30601	Hepatitis
34	167	100.0	2201	5	ABG30584	Hepatitis
35	167	100.0	2201	5	ABG30591	Hepatitis
36	167	100.0	2201	5	ABG30600	Hepatitis
37	167	100.0	2201	5	ABG30581	Hepatitis
38	167	100.0	2201	5	ABG30586	Hepatitis
39	167	100.0	2201	5	ABG30593	Hepatitis
40	167	100.0	2201	5	ABG30582	Hepatitis
41	167	100.0	2201	5	ABG30580	Hepatitis
42	167	100.0	2201	5	ABG30602	Hepatitis
43	167	100.0	2201	5	ABG30587	Hepatitis
44	167	100.0	2201	5	ABG30589	Hepatitis
45	167	100.0	2201	5	ABG30599	Hepatitis

```

XX      Sequence 33 AA;
SQ
      Query Match      100.0%; Score 167; DB 4; Length 33;
      Best Local Similarity 100.0%; Pred. No. 1e-19;
      Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MNRLIAPASRGNHVSPHYVPESDAAARVTQIL 33
      |||||
Db      1 MNRLIAPASRGNHVSPHYVPESDAAARVTQIL 33

RESULT 2
ID      AAR25865 standard; protein; 196 AA.
XX
AC      AAR25865;
XX
DT      21-JAN-1993 (first entry)
XX
DE      HCV polypeptide 12.
XX
KW      Recombinant vector; E. coli; diagnostic; reagent; type C hepatitis.
XX
OS      Hepatitis C virus.
XX
PN      JP04179482-A.
XX
PD      26-JUN-1992.
XX
PF      11-NOV-1990; 90JP-00304417.
XX
PR      11-NOV-1990; 90JP-00304417.
XX
PA      (TOKU ) TOKUYAMA SODA KK.
XX
DR      WPI; 1992-263663/32.
XX
N-PSDB; AAQ26992.
XX
Hepatitis C virus antigen expressed as recombinant in E.coli - useful for
diagnosis of hepatitis C virus infection.
XX
Claim 1; Page 4; 66pp; Japanese.
XX
The sequences given in AAR25854-74 are hepatitis C virus protiens. The
genes encoding these proteins can each be used to prepare recombinant
vectors by ligating the gene of interest in to a vector to be expressed
in E. coli. These polypeptides are useful as diagnostic reagents for type
C hepatitis and they may be produced efficiently by recombinant methods
XX
SQ      Sequence 196 AA;

      Query Match      100.0%; Score 167; DB 2; Length 196;
      Best Local Similarity 100.0%; Pred. No. 9.8e-19;
      Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MNRLIAPASRGNHVSPHYVPESDAAARVTQIL 33
      |||||
Db      6 MNRLIAPASRGNHVSPHYVPESDAAARVTQIL 38

RESULT 3
ID      AAW41743 standard; protein; 196 AA.
XX
AC      AAW41743;
XX
DT      22-MAY-1998 (first entry)
XX
DE      Hepatitis C virus antigen.
XX
KW      non-A non-B hepatitis virus; NANBH; hepatitis C virus; HCV; antigen;
diagnosis; detection.

```

```

XX      Hepatitis virus.
OS
XX      JP05176774-A.
PN
XX      20-JUL-1993.
PD
XX      18-DEC-1991; 91JP-00354708.
PF
XX      18-DEC-1990; 90JP-00412020.
PR
XX      (SHIM/) SHIMOTONO K.
PA      (GREC ) GREEN CROSS CORP.
XX
DR      WPI; 1993-260858/33.
XX
Protein contg. non-A non-B hepatitis antigen fragment - prepd. by
culturing transformants transformed by vector contg. base sequence coding
specified aminoacid sequences, used for detecting hepatitis.
XX
Claim 1; Fig 9; 53pp; Japanese.
PS
XX
The present sequence is a non-A non-B hepatitis virus (NANBH) or
hepatitis C virus (HCV) antigen, useful for diagnosis or detection
XX
SQ      Sequence 196 AA;

      Query Match      100.0%; Score 167; DB 2; Length 196;
      Best Local Similarity 100.0%; Pred. No. 9.8e-19;
      Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MNRLIAPASRGNHVSPHYVPESDAAARVTQIL 33
      |||||
Db      6 MNRLIAPASRGNHVSPHYVPESDAAARVTQIL 38

RESULT 4
ID      ADO36214 standard; protein; 214 AA.
XX
AC      ADO36214;
XX
DT      26-AUG-2004 (first entry)
XX
DE      Hepatitis C virus (HCV) truncated NS4B protein.
XX
KW      hepatotropic; virucide; vaccine; gene therapy; vaccine;
Hepatitis C virus; HCV; core protein; HCV infection; vaccination; NS4B.
XX
OS      Hepatitis C virus.
XX
PN      WO2004046175-A1.
XX
PD      03-JUN-2004.
XX
PF      13-NOV-2003; 2003WO-EP012793.
XX
PR      15-NOV-2002; 2002GB-00026722.
XX
PA      (GLAX ) GLAXO GROUP LTD.
XX
PI      Brett S, Hamblin PA, Ogilvie L;
XX
DR      WPI; 2004-420613/39.
XX
N-PSDB; ADO36225.
XX
New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
the polypeptide sequences of the HCV core and at least one other HCV
protein, for use in medicine, particularly for manufacturing a medicament
for treating HCV.
XX
Example 2; Page 27; 78pp; English.
PS
XX

```

CC The invention describes a polynucleotide vaccine comprising a
 CC polynucleotide sequence (S1) encoding the Hepatitis C virus (HCV) Core
 CC protein and at least 1 other HCV protein, and causes expression of the
 CC proteins in cells (in which (S1) has been mutated or positioned relative
 CC to the polynucleotide sequence encoding the other HCV protein, so that
 CC the negative effect of the Core protein on expression of the other HCV
 CC protein is reduced). Also described are: a method of preventing or
 CC treating an HCV infection in a mammal, comprising administering the
 CC vaccine cited above to a mammal; and a method of vaccination of an
 CC individual, comprising taking a polynucleotide vaccine as cited above,
 CC coating the polynucleotide onto gold beads and delivering the gold beads
 CC into the skin. HCV nucleic acids, polypeptides, host cells, vectors and
 CC antibodies used in the methods, are also disclosed. The polynucleotide
 CC vaccine is useful in the manufacture of a medicament for the treatment of
 CC HCV. This is the amino acid sequence of the truncated HCV NS4B protein.

XX Sequence 214 AA;

Query Match 100.0%; Score 167; DB 8; Length 214;
 Best Local Similarity 100.0%; Pred. No. 1.1e-18;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLIAPASRGNHVSPTHYPESDAAARVTQIL 33
 |||||
 Db 158 MNRLIAPASRGNHVSPTHYPESDAAARVTQIL 190

RESULT 5

ADO79388
 ID ADO79388 standard; protein; 214 AA.

AC ADO79388;

XX 26-AUG-2004 (first entry)

XX Hepatitis C virus NS4B protein.

XX HCV; NS4A; vaccine; DNA immunisation; hepatotropic; virucide; mutant;
 KW mutin.

XX Hepatitis C virus.

OS Synthetic.

XX WO2004046176-A1.

XX 03-JUN-2004.

XX 13-NOV-2003; 2003WO-EP012830.

XX 15-NOV-2002; 2002GB-00026722.

XX (GLAX) GLAXO GROUP LTD.

XX Brett S, Hamblin PA, Ogilvie L;

XX WPI; 2004-420614/39.

XX N-PSDB; ADO79399.

XX New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
 PT the polypeptide sequences of the HCV core, NS3, NS4B and NS5B proteins,
 PT for use in medicine, in particular for manufacturing a medicament for the
 PT treatment of HCV.

XX Example 2; Page 27; 79pp; English.

XX The present sequence is that of the NS4A protein of hepatitis C virus
 CC (HCV). The sequence is the translation sequence of a polynucleotide in
 CC which codon usage was altered to resemble that of highly expressed human
 CC genes. HCV vaccines of the invention comprise a polynucleotide that
 CC encodes the HCV proteins Core, NS3, NS4B and NS5B, and does not encode
 CC the NS4A and/or NS5A proteins. The proteins may be expressed as
 CC individual proteins or as fusion proteins. Preferred fusions include
 CC double fusions between NS4B and NS5B and between Core and NS3. The

CC vaccines are useful for the treatment or prevention of an HCV infection.

XX Sequence 214 AA;

Query Match 100.0%; Score 167; DB 8; Length 214;

Best Local Similarity 100.0%; Pred. No. 1.1e-18;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLIAPASRGNHVSPTHYPESDAAARVTQIL 33

Db 158 MNRLIAPASRGNHVSPTHYPESDAAARVTQIL 190

RESULT 6

AAR25887

ID AAR25887 standard; protein; 219 AA.

XX AAR25887;

XX 09-SEP-2004 (revised)

DT 21-JAN-1993 (first entry)

XX HK12.

XX Recombinant vector; E. coli; diagnostic; reagent; type C hepatitis.

XX Hepatitis C virus.

OS Unidentified.

XX Key Location/Qualifiers

FT Protein 11.206

FT /note= "Sequence AAR25865"

XX JP04179482-A.

XX 26-JUN-1992.

XX 11-NOV-1990; 90JP-00304417.

XX 11-NOV-1990; 90JP-00304417.

XX (TOKU) TOKUYAMA SODA KK.

XX WPI; 1992-263663/32.

XX N-PSDB; AAQ27014.

XX Hepatitis C virus antigen expressed as recombinant in E.coli - useful for
 PT diagnosis of hepatitis C virus infection.

XX Disclosure; Fig 13; 66pp; Japanese.

XX The sequences given in AAR25876-95 are encoded by the claimed hepatitis C
 CC virus genes of the invention which have been inserted into an E. coli
 CC vector. These polypeptides are useful as diagnostic reagents for type C
 CC hepatitis and they may be produced efficiently by recombinant DNA
 CC techniques

XX Revised record issued on 09-SEP-2004 : Correction to feature table key

XX Sequence 219 AA;

Query Match 100.0%; Score 167; DB 2; Length 219;

Best Local Similarity 100.0%; Pred. No. 1.1e-18;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLIAPASRGNHVSPTHYPESDAAARVTQIL 33

Db 16 MNRLIAPASRGNHVSPTHYPESDAAARVTQIL 48

RESULT 7

AAR29882

ID AAR29882 standard; protein; 272 AA.

KW turn structure; alpha helix; beta sheet; antigen; determinant; antiserum.
 XX Hepatitis C virus.
 OS
 XX EP518313-A2.
 PN
 XX 16-DEC-1992.
 PD
 XX 11-JUN-1992; 92EP-00109812.
 PF
 XX 11-JUN-1991; 91JP-00139268.
 PR 12-JUL-1991; 91JP-00172794.
 PR 07-OCT-1991; 91JP-00287008.
 PR 16-DEC-1991; 91JP-00332329.
 PR 20-APR-1992; 92JP-00099957.
 XX
 XX (MITU) MITSUBISHI KASEI CORP.
 PA
 XX Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;
 PI
 XX WPI: 1992-417213/51.
 DR N-PSDB; AAQ32516.
 DR
 XX New hepatitis C virus gene and its encoded protein - used for diagnosing
 PT and vaccinating against hepatitis C virus infections.
 PT
 XX Disclosure; Page 229-31; 305pp; English.
 PS
 XX The sequences given in AAR29871-906 and AAR29533 are encoded by various
 CC clones of the NS4-NS5 regions of the Hepatitis C Virus (HCV) gene of the
 CC invention. These NS4-NS5 RNA sequences were isolated from the serum of a
 CC patient suffering from hepatitis C (HC). The isolated RNA sequences were
 CC converted into cDNA using transcriptase in the presence of one of the
 CC primer sequences given in AAQ32563-77. The sequences were then amplified
 CC using primer pairs. The cDNA sequences isolated represent different
 CC alleles of the same region of the HCV gene. Sequence analysis shows that
 CC these clones represent the core region and some upstream sequences of
 CC HCV. These polypeptides are thought to contain a highly hydrophilic
 CC region which can adopt a "turn structure" which is not an alpha helix or
 CC a beta sheet. These polypeptides are thought to act as antigen
 CC determinants and are highly reactive with antiserum raised against HCV-
 CC associated antigens. See also AAQ32436. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX
 XX Sequence 272 AA;
 SQ
 Query Match 100.0%; Score 167; DB 2; Length 272;
 Best Local Similarity 100.0%; Pred. No. 1.5e-18;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 MNRLLAFASRGHVSPTHVPESDAARVTOIL 33
 DB 12 MNRLLAFASRGHVSPTHVPESDAARVTOIL 44
 RESULT 10
 AAR29871
 ID AAR29871 standard; protein; 272 AA.
 XX
 XX AAR29871;
 AC
 XX
 XX 25-MAR-2003 (revised)
 DT 26-APR-1993 (first entry)
 DT
 XX HCV NS4-NS5 peptide N22-1, N22-3, H22-8, H22-9.
 DE
 XX Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;
 KW transcriptase; cDNA; primer; allele; core; region; upstream; hydrophilic;
 KW turn structure; alpha helix; beta sheet; antigen; determinant; antiserum.
 XX
 OS Hepatitis C virus.
 XX
 XX Key Location/Qualifiers
 PH

FT Misc-difference 46 /label= Asn, Ser
 FT Misc-difference 62 /label= Glu, Asp
 FT Misc-difference 85 /label= Ala, Ser
 FT Misc-difference 87 /label= Cys, Phe
 FT Misc-difference 105 /label= Phe, Leu
 FT Misc-difference 122 /label= Tyr, Gln, His
 FT Misc-difference 132 /label= Thr, Ala
 FT Misc-difference 143 /label= Val, Thr
 FT Misc-difference 145 /label= Pro, Leu
 FT Misc-difference 153 /label= His, Arg
 FT Misc-difference 176 /label= Arg, Lys
 FT Misc-difference 183 /label= Ile, Ala
 FT Misc-difference 206 /label= Val, Leu
 FT Misc-difference 227 /label= His, Arg
 FT Misc-difference 250 /label= Tyr, Phe
 FT Misc-difference 251 /label= Thr, Pro
 FT Misc-difference 265 /label= Thr, Met, Ala
 XX
 XX EP518313-A2.
 PN
 XX 16-DEC-1992; 92EP-00109812.
 PD
 XX 11-JUN-1992; 91JP-00139268.
 PR 12-JUL-1991; 91JP-00172794.
 PR 07-OCT-1991; 91JP-00287008.
 PR 16-DEC-1991; 91JP-00332329.
 PR 20-APR-1992; 92JP-00099957.
 XX
 XX (MITU) MITSUBISHI KASEI CORP.
 PA
 XX Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;
 PI
 XX WPI: 1992-417213/51.
 DR N-PSDB; AAQ32502.
 DR
 XX New hepatitis C virus gene and its encoded protein - used for diagnosing
 PT and vaccinating against hepatitis C virus infections.
 PT
 XX Disclosure; Page 193-94; 305pp; English.
 PS
 XX The sequences given in AAR29871-906 and AAR29533 are encoded by various
 CC clones of the NS4-NS5 regions of the Hepatitis C Virus (HCV) gene of the
 CC invention. These NS4-NS5 RNA sequences were isolated from the serum of a
 CC patient suffering from hepatitis C (HC). The isolated RNA sequences were
 CC converted into cDNA using transcriptase in the presence of one of the
 CC primer sequences given in AAQ32563-77. The sequences were then amplified
 CC using primer pairs. The cDNA sequences isolated represent different
 CC alleles of the same region of the HCV gene. Sequence analysis shows that
 CC these clones represent the core region and some upstream sequences of
 CC HCV. These polypeptides are thought to contain a highly hydrophilic
 CC region which can adopt a "turn structure" which is not an alpha helix or
 CC a beta sheet. These polypeptides are thought to act as antigen
 CC determinants and are highly reactive with antiserum raised against HCV-
 CC associated antigens. See also AAQ32436. (Updated on 25-MAR-2003 to

```

CC correct PN field.)
XX Best Local Similarity 100.0%; Score 167; DB 2; Length 272;
SQ Sequence 272 AA; Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match
Best Local Similarity 100.0%; Pred. No. 1.5e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLIAPASRGNHVSPTHYVPESDAARVTQIL 33
Db 12 MNRLIAPASRGNHVSPTHYVPESDAARVTQIL 44

RESULT 11
AAR29884
ID AAR29884 standard; protein; 272 AA.
XX AAR29884;
XX 25-MAR-2003 (revised)
DT 26-APR-1993 (first entry)
XX HCV NS4-NS5 peptide H22-3.
XX Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;
KW transcriptase; cDNA; primer; allele; core; region; upstream; hydrophilic;
KW turn structure; alpha helix; beta sheet; antigen; determinant; antiserum.
XX Hepatitis C virus.
XX EP518313-A2.
XX 16-DEC-1992.
XX 11-JUN-1992; 92EP-00109812.
XX 11-JUN-1991; 91JP-00139268.
PR 12-JUL-1991; 91JP-00172794.
PR 07-OCT-1991; 91JP-00287008.
PR 16-DEC-1991; 91JP-00332329.
XX 20-APR-1992; 92JP-00099957.
XX (MITU ) MITSUBISHI KASEI CORP.
XX Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;
WPI; 1992-417213/51.
DR N-PSDB; AAQ32515.
XX New hepatitis C virus gene and its encoded protein - used for diagnosing
PT and vaccinating against hepatitis C virus infections.
XX Disclosure; Page 227-29; 305pp; English.
XX The sequences given in AAR29871-906 and AAR29533 are encoded by various
CC clones of the NS4-NS5 regions of the Hepatitis C Virus (HCV) gene of the
CC invention. These NS4-NS5 RNA sequences were isolated from the serum of a
CC patient suffering from hepatitis C (HC). The isolated RNA sequences were
CC converted into cDNA using transcriptase in the presence of one of the
CC primer sequences given in AAQ32565-77. The sequences were then amplified
CC using primer pairs. The cDNA sequences isolated represent different
CC alleles of the same region of the HCV gene. Sequence analysis shows that
CC these clones represent the core region and some upstream sequences of
CC HCV. These polypeptides are thought to contain a highly hydrophilic
CC region which can adopt a "turn structure" which is not an alpha helix or
CC a beta sheet. These polypeptides are thought to act as antigen
CC determinants and are highly reactive with antiserum raised against HCV-
CC associated antigens. See also AAQ32436. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX Sequence 272 AA;
SQ Query Match 100.0%; Score 167; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.5e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLIAPASRGNHVSPTHYVPESDAARVTQIL 33
Db 12 MNRLIAPASRGNHVSPTHYVPESDAARVTQIL 44

RESULT 12
AAR29877
ID AAR29877 standard; protein; 360 AA.
XX AAR29877;
XX 25-MAR-2003 (revised)
DT 26-APR-1993 (first entry)
XX HCV NS4-NS5 peptide 2217.
XX Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;
KW transcriptase; cDNA; primer; allele; core; region; upstream; hydrophilic;
KW turn structure; alpha helix; beta sheet; antigen; determinant; antiserum.
XX Hepatitis C virus.
XX EP518313-A2.
XX 16-DEC-1992.
XX 11-JUN-1992; 92EP-00109812.
XX 11-JUN-1991; 91JP-00139268.
PR 12-JUL-1991; 91JP-00172794.
PR 07-OCT-1991; 91JP-00287008.
PR 16-DEC-1991; 91JP-00332329.
XX 20-APR-1992; 92JP-00099957.
XX (MITU ) MITSUBISHI KASEI CORP.
XX Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;
WPI; 1992-417213/51.
DR N-PSDB; AAQ32508.
XX New hepatitis C virus gene and its encoded protein - used for diagnosing
PT and vaccinating against hepatitis C virus infections.
XX Disclosure; Page 202-04; 305pp; English.
XX The sequences given in AAR29871-906 and AAR29533 are encoded by various
CC clones of the NS4-NS5 regions of the Hepatitis C Virus (HCV) gene of the
CC invention. These NS4-NS5 RNA sequences were isolated from the serum of a
CC patient suffering from hepatitis C (HC). The isolated RNA sequences were
CC converted into cDNA using transcriptase in the presence of one of the
CC primer sequences given in AAQ32565-77. The sequences were then amplified
CC using primer pairs. The cDNA sequences isolated represent different
CC alleles of the same region of the HCV gene. Sequence analysis shows that
CC these clones represent the core region and some upstream sequences of
CC HCV. These polypeptides are thought to contain a highly hydrophilic
CC region which can adopt a "turn structure" which is not an alpha helix or
CC a beta sheet. These polypeptides are thought to act as antigen
CC determinants and are highly reactive with antiserum raised against HCV-
CC associated antigens. See also AAQ32436. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX Sequence 360 AA;
SQ Query Match 100.0%; Score 167; DB 2; Length 360;
Best Local Similarity 100.0%; Pred. No. 2.1e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLIAPASRGNHVSPTHYVPESDAARVTQIL 33
Db 12 MNRLIAPASRGNHVSPTHYVPESDAARVTQIL 44

```

```

Db      12 MNRLIAPASRGNHVSPTHVPESDAAARVTQIL 44
RESULT 13
AAR80044
ID      AAR80044 standard; protein; 767 AA.
XX
AC      AAR80044;
XX
DT      25-MAR-1996 (first entry)
XX
DE      Hepatitis C virus antigenic sequence.
XX
KW      Maltose binding protein; MBP; non-A non-B; HCV; diagnosis; antigen.
XX
OS      Hepatitis C virus.
XX
PN      JP07198723-A.
XX
PD      01-AUG-1995.
XX
PF      29-DEC-1993; 93JP-00351227.
XX
PR      29-DEC-1993; 93JP-00351227.
XX
PA      (JAPG ) NIPPON ZEON KK.
XX
DR      WPI; 1995-300583/39.
XX
N-PSDB; AAT04565.
XX
Diagnosis of HCV infection - using a fused protein comprising a HCV
PT      antigen and a carrier protein.
XX
PS      Example 2; Page 7-10; 10pp; Japanese.
XX
AAT04565 encodes AAR80044 a hepatitis C virus (HCV) antigenic sequence.
CC      An antigen derived from the antigenic sequence was combined with a
CC      carrier protein (maltose binding protein) to produce a fusion protein,
CC      which can be used for the highly sensitive detection of HCV
XX
SQ      Sequence 767 AA;
Query Match      100.0%; Score 167; DB 2; Length 767;
Best Local Similarity 100.0%; Pred. No. 5.6e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 MNRLIAPASRGNHVSPTHVPESDAAARVTQIL 33
DB      732 MNRLIAPASRGNHVSPTHVPESDAAARVTQIL 764
RESULT 14
AAR29881
ID      AAR29881 standard; protein; 863 AA.
XX
AC      AAR29881;
XX
DT      25-MAR-2003 (revised)
DT      26-APR-1993 (first entry)
XX
DE      HCV NS4-NS5 peptide 2218.
XX
KW      Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;
KW      transcriptase; cDNA; primer; allele; core; region; upstream; hydrophilic;
KW      turn structure; alpha helix; beta sheet; antigen; determinant; antiserum.
XX
OS      Hepatitis C virus.
XX
FH      Key Location/Qualifiers
FT      Misc-difference 583
FT      /label= Leu
XX
EP518313-A2.
XX
16-DEC-1992.
XX
11-JUN-1992; 92EP-00109812.
XX
11-JUN-1991; 91JP-00139268.
PR      12-JUL-1991; 91JP-00172794.
PR      07-OCT-1991; 91JP-00287008.
PR      16-DEC-1991; 91JP-00332329.
PR      20-APR-1992; 92JP-00099957.
XX
(MITU ) MITSUBISHI KASEI CORP.
XX
Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;
WPI; 1992-417213/51.
N-PSDB; AAQ32512.
XX
New hepatitis C virus gene and its encoded protein - used for diagnosing
PT      and vaccinating against hepatitis C virus infections.
XX
Disclosure; Page 212-16; 305pp; English.
XX
The sequences given in AAR29871-906 and AAR29533 are encoded by various
CC      clones of the NS4-NS5 regions of the Hepatitis C Virus (HCV) gene of the
CC      invention. These NS4-NS5 RNA sequences were isolated from the serum of a
CC      patient suffering from hepatitis C (HC). The isolated RNA sequences were
CC      converted into cDNA using transcriptase in the presence of one of the
CC      primer sequences given in AAQ32565-77. The sequences were then amplified
CC      using primer pairs. The cDNA sequences isolated represent different
CC      alleles of the same region of the HCV gene. Sequence analysis shows that
CC      these clones represent the core region and some upstream sequences of
CC      HCV. These polypeptides are thought to contain a highly hydrophilic
CC      region which can adopt a "turn structure" which is not an alpha helix or
CC      a beta sheet. These polypeptides are thought to act as antigen
CC      determinants and are highly reactive with antiserum raised against HCV-
CC      associated antigens. See also AAQ32436. (Updated on 25-MAR-2003 to
CC      correct PN field.)
XX
SQ      Sequence 863 AA;
Query Match      100.0%; Score 167; DB 2; Length 863;
Best Local Similarity 100.0%; Pred. No. 6.5e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 MNRLIAPASRGNHVSPTHVPESDAAARVTQIL 33
DB      12 MNRLIAPASRGNHVSPTHVPESDAAARVTQIL 44
RESULT 15
AAR12599
ID      AAR12599 standard; protein; 1250 AA.
XX
AC      AAR12599;
XX
DT      25-MAR-2003 (revised)
DT      17-SEP-1991 (first entry)
XX
DE      Portion of PT-NANBH viral non-structural protein.
XX
KW      post-transfusional non-A, non-B hepatitis; virus; vaccine; ss.
XX
OS      Non-A.
OS      non-B hepatitis virus.
XX
PN      GB2239245-A.
XX
PD      26-JUN-1991.
XX
PF      17-DEC-1990; 90GB-00027250.
XX
PR      18-DEC-1989; 89GB-00028562.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:04:54 ; Search time 9.24324 Seconds
(without alignments)
266.510 Million cell updates/sec

Title: US-09-758-308-4

Perfect score: 167

Sequence: 1 MRLIAFASRGHVSPTHYVPESDAARVTQIL 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pap:*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pap:*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pap:*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pap:*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pap:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	167	100.0	151	1	US-08-324-977-48
2	167	100.0	151	2	US-08-384-616-48
3	167	100.0	151	2	US-08-904-686A-48
4	167	100.0	151	3	US-09-315-850-48
5	167	100.0	1985	4	US-09-539-601-9
6	167	100.0	1985	4	US-09-539-601-12
7	167	100.0	1985	4	US-09-539-601-24
8	167	100.0	1985	4	US-09-539-601-30
9	167	100.0	2013	1	US-08-324-977-12
10	167	100.0	2013	2	US-08-384-616-12
11	167	100.0	2013	2	US-08-904-686A-12
12	167	100.0	22013	3	US-09-315-850-12
13	167	100.0	22013	3	US-08-952-981A-2
14	167	100.0	2201	4	US-09-539-601-6
15	167	100.0	2201	4	US-09-539-601-15
16	167	100.0	2201	4	US-10-029-907-3
17	167	100.0	2620	1	US-08-324-977-32
18	167	100.0	2620	2	US-08-384-616-32
19	167	100.0	2620	2	US-08-904-686A-32
20	167	100.0	2620	3	US-09-315-850-32
21	167	100.0	2621	1	US-08-324-977-36
22	167	100.0	2621	2	US-08-384-616-36
23	167	100.0	2621	2	US-08-904-686A-36
24	167	100.0	2621	3	US-09-315-850-36
25	167	100.0	3010	1	US-08-324-977-2
26	167	100.0	3010	1	US-08-324-977-14
27	167	100.0	3010	2	US-08-384-616-2

28	167	100.0	3010	2	US-08-384-616-14	Sequence 14, Appl
29	167	100.0	3010	2	US-08-904-686A-2	Sequence 2, Appl
30	167	100.0	3010	2	US-08-904-686A-14	Sequence 14, Appl
31	167	100.0	3010	3	US-09-014-416-3	Sequence 3, Appl
32	167	100.0	3010	3	US-09-315-850-2	Sequence 2, Appl
33	167	100.0	3010	3	US-09-315-850-14	Sequence 14, Appl
34	167	100.0	3010	4	US-09-539-601-3	Sequence 3, Appl
35	167	100.0	3010	4	US-09-539-601-27	Sequence 27, Appl
36	167	100.0	3010	4	US-09-539-601-33	Sequence 33, Appl
37	161	96.4	106	3	US-08-444-818-24	Sequence 24, Appl
38	161	96.4	260	4	US-10-104-966-8	Sequence 8, Appl
39	161	96.4	269	3	US-09-100-557-1	Sequence 1, Appl
40	161	96.4	360	3	US-08-850-328-4	Sequence 4, Appl
41	161	96.4	859	3	US-08-444-818-30	Sequence 30, Appl
42	161	96.4	1786	3	US-08-444-818-54	Sequence 54, Appl
43	161	96.4	2261	3	US-08-444-818-66	Sequence 66, Appl
44	161	96.4	2436	3	US-08-444-818-75	Sequence 75, Appl
45	161	96.4	2772	3	US-08-444-818-89	Sequence 89, Appl

ALIGNMENTS

RESULT 1
US-08-324-977-48
; Sequence 48, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeand &
; ADDRESS: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281

; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-324-977-48

Query Match 100.0%; Score 167; DB 1; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.3e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLIAPASGNHVSPTHYPESDAARVTQIL 33
|||||
Db 54 MNRLIAPASGNHVSPTHYPESDAARVTQIL 86

RESULT 2

US-08-384-616-48
; Sequence 48, Application US/08384616
; Patent No. 5847101

; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357

; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-384-616-48

Query Match 100.0%; Score 167; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.3e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLIAPASGNHVSPTHYPESDAARVTQIL 33
|||||
Db 54 MNRLIAPASGNHVSPTHYPESDAARVTQIL 86

RESULT 3

US-08-904-686A-48
; Sequence 48, Application US/08904686A
; Patent No. 5998130

; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeLeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 151 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-904-686A-48

Query Match 100.0%; Score 167; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.3e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNRLIAFASRGHVSPTHYVPESDAAARVTQIL 33
Db 54 MNRLIAFASRGHVSPTHYVPESDAAARVTQIL 86

RESULT 4

US-09-315-850-48
Sequence 48, Application US/09315850
Patent No. 6217872

GENERAL INFORMATION:

APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKU, Ieao
APPLICANT: MORI, Chisato
APPLICANT: TANAKIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, McLeLland &
ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/315.850
FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/904,686
FILING DATE: 01-AUG-1997

APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990

ATTORNEY/AGENT INFORMATION:

NAME: McLeLland, Le-Nhung

REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 151 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-315-850-48

Query Match 100.0%; Score 167; DB 3; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.3e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNRLIAFASRGHVSPTHYVPESDAAARVTQIL 33
Db 54 MNRLIAFASRGHVSPTHYVPESDAAARVTQIL 86

RESULT 5

US-09-539-601-9
Sequence 9, Application US/09539601C
Patent No. 6630343

GENERAL INFORMATION:

APPLICANT: Bartenschlager, Ralf FW
TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
FILE REFERENCE: all sequences
CURRENT APPLICATION NUMBER: US/09/539,601C
CURRENT FILING DATE: 2001-08-30
EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
EARLIER FILING DATE: 1999-04-03
NUMBER OF SEQ ID NOS: 51
SOFTWARE: Patentin ver. 2.1
SEQ ID NO 9

LENGTH: 1985

TYPE: PRT

ORGANISM: Hepatitis C virus

US-09-539-601-9

Query Match 100.0%; Score 167; DB 4; Length 1985;
Best Local Similarity 100.0%; Pred. No. 3.3e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNRLIAFASRGHVSPTHYVPESDAAARVTQIL 33
Db 891 MNRLIAFASRGHVSPTHYVPESDAAARVTQIL 923

RESULT 6

US-09-539-601-12
Sequence 12, Application US/09539601C
Patent No. 6630343

GENERAL INFORMATION:

APPLICANT: Bartenschlager, Ralf FW
TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
FILE REFERENCE: all sequences
CURRENT APPLICATION NUMBER: US/09/539,601C
CURRENT FILING DATE: 2001-08-30
EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
EARLIER FILING DATE: 1999-04-03
NUMBER OF SEQ ID NOS: 51
SOFTWARE: Patentin ver. 2.1
SEQ ID NO 12

LENGTH: 1985

TYPE: PRT

ORGANISM: Hepatitis C virus

US-09-539-601-12

Query Match 100.0%; Score 167; DB 4; Length 1985;
Best Local Similarity 100.0%; Pred. No. 3.3e-17;

APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/384,616
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703B
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2013 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-384-616-12

Query Match 100.0%; Score 167; DB 2; Length 2013;
Best Local Similarity 100.0%; Pred. No. 3.4e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLIAPASRGHNVSPHYVPESDAAARVTQIL 33
DB 1916 MNRLIAPASRGHNVSPHYVPESDAAARVTQIL 1948

RESULT 11
US-08-904-686A-12
Sequence 12, Application US/08904686A
Patent No. 5998130
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKU, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/904,686A
FILING DATE: 01-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: McLeland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2013 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-904-686A-12

Query Match 100.0%; Score 167; DB 2; Length 2013;
Best Local Similarity 100.0%; Pred. No. 3.4e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLIAPASRGHNVSPHYVPESDAAARVTQIL 33
DB 1916 MNRLIAPASRGHNVSPHYVPESDAAARVTQIL 1948

RESULT 12
US-09-315-850-12
Sequence 12, Application US/09315850
Patent No. 6217872
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKU, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
CDNA AND ANTIGEN POLYPEPTIDE

;; NUMBER OF SEQUENCES: 50
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLand &
;; ADDRESSEE: Naughton
;; STREET: 1725 K St. N.W. Suite 1000
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: U.S.A.
;; ZIP: 20006

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
;; SOFTWARE: ASCII

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/315.850
;; FILING DATE:

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/904.686
;; FILING DATE: 01-AUG-1997
;; APPLICATION NUMBER: US 08/324.977
;; FILING DATE: 18-OCT-1994

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: JP 2-167466

;; FILING DATE: 25-JUN-1990

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: JP 2-230921

;; FILING DATE: 31-AUG-1990

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: JP 2-305605

;; FILING DATE: 09-NOV-1990

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 08/099.706

;; FILING DATE: 30-JUL-1993

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 07/769.996

;; FILING DATE: 02-OCT-1991

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 07/635.451

;; FILING DATE: 28-DEC-1990

;; ATTORNEY/AGENT INFORMATION:

;; NAME: McLeLand, Le-Nhung

;; REGISTRATION NUMBER: 31,541

;; REFERENCE/DOCKET NUMBER: 900703G

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (202) 659-2930

;; TELEFAX: (202) 887-0357

;; INFORMATION FOR SEQ ID NO: 12:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 2013 amino acids

;; TYPE: amino acid

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

;; US-09-315-850-12

Query Match 100.0%; Score 167; DB 3; Length 2013;
Best Local Similarity 100.0%; Pred. No. 3.4e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MNRLIAFASRGNHVSPTHYPESDAAARVTQIL 33
|||||

Db 1916 MNRLIAFASRGNHVSPTHYPESDAAARVTQIL 1948

RESULT 13

US-08-952-981A-2

;; Sequence 2, Application US/08952981A

;; Patent No. 6383768

;; GENERAL INFORMATION:

;; APPLICANT: DE FRANCESCO, Raffaele

;; APPLICANT: TOMEI, Licia

;; APPLICANT: BEHRENS, Sven-Erik

;; TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE RNA-DEPENDENT RNA

;; TITLE OF INVENTION: POLYMERASE AND TERMINAL NUCLEOTIDYL TRANSFERASE
;; TITLE OF INVENTION: ACTIVITIES ENCODED BY HEPATITIS C VIRUS (HCV)
;; FILE REFERENCE: IT0002P
;; CURRENT APPLICATION NUMBER: US/08/952.981A
;; CURRENT FILING DATE: 1998-03-23
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 2
;; LENGTH: 2201
;; TYPE: PRT
;; ORGANISM: CDNA clone pCD (38-9.4)
;; US-08-952-981A-2

Query Match 100.0%; Score 167; DB 3; Length 2201;
Best Local Similarity 100.0%; Pred. No. 3.8e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MNRLIAFASRGNHVSPTHYPESDAAARVTQIL 33
|||||

Db 1107 MNRLIAFASRGNHVSPTHYPESDAAARVTQIL 1139

RESULT 14

US-09-539-601-6

;; Sequence 6, Application US/09539601C

;; Patent No. 6630343

;; GENERAL INFORMATION:

;; APPLICANT: Bartenschlager, Ralf FW

;; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System

;; FILE REFERENCE: all sequences

;; CURRENT APPLICATION NUMBER: US/09/539.601C

;; CURRENT FILING DATE: 2001-08-30

;; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY

;; EARLIER FILING DATE: 1999-04-03

;; NUMBER OF SEQ ID NOS: 51

;; SOFTWARE: PatentIn Ver. 2.1

;; SEQ ID NO 6

;; LENGTH: 2201

;; TYPE: PRT

;; ORGANISM: Hepatitis C virus

;; US-09-539-601-6

Query Match 100.0%; Score 167; DB 4; Length 2201;
Best Local Similarity 100.0%; Pred. No. 3.8e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MNRLIAFASRGNHVSPTHYPESDAAARVTQIL 33
|||||

Db 1107 MNRLIAFASRGNHVSPTHYPESDAAARVTQIL 1139

RESULT 15

US-09-539-601-15

;; Sequence 15, Application US/09539601C

;; Patent No. 6630343

;; GENERAL INFORMATION:

;; APPLICANT: Bartenschlager, Ralf FW

;; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System

;; FILE REFERENCE: all sequences

;; CURRENT APPLICATION NUMBER: US/09/539.601C

;; CURRENT FILING DATE: 2001-08-30

;; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY

;; EARLIER FILING DATE: 1999-04-03

;; NUMBER OF SEQ ID NOS: 51

;; SOFTWARE: PatentIn Ver. 2.1

;; SEQ ID NO 15

;; LENGTH: 2201

;; TYPE: PRT

;; ORGANISM: Hepatitis C virus

;; US-09-539-601-15

Query Match 100.0%; Score 167; DB 4; Length 2201;
Best Local Similarity 100.0%; Pred. No. 3.8e-17;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MNRLIAFASRGNHVSPTHYVPESDAAARVTOIL 33
Db 1107 MNRLIAFASRGNHVSPTHYVPESDAAARVTOIL 1139

Search completed: August 12, 2005, 14:24:49
Job time : 10.2432 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:03:18 ; Search time 30.4865 Seconds
(without alignments)
554.298 Million cell updates/sec

Title: US-09-758-308-4

Perfect score: 167

Sequence: 1 MNRLIAPASGNHVSPTHYVPESDAARVTQIL 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	167	100.0	1805	2	O41809 hepatitis c
2	167	100.0	1984	2	Q74V8 hepatitis c
3	167	100.0	2284	2	O81817 hepatitis c
4	167	100.0	2864	2	O92973 hepatitis c
5	167	100.0	2864	2	O92974 hepatitis c
6	167	100.0	2864	2	O92975 hepatitis c
7	167	100.0	2864	2	O92976 hepatitis c
8	167	100.0	2864	2	O92977 hepatitis c
9	167	100.0	2864	2	O92978 hepatitis c
10	167	100.0	2864	2	O92979 hepatitis c
11	167	100.0	3008	2	O92980 hepatitis c
12	167	100.0	3010	1	POLG HCVBK
13	167	100.0	3010	1	POLG HCVUA
14	167	100.0	3010	1	POLG HCVUT
15	167	100.0	3010	1	POLG HCVTV
16	167	100.0	3010	2	O92979 hepatitis c
17	167	100.0	3010	2	O92979 hepatitis c
18	167	100.0	3010	2	O92979 hepatitis c
19	167	100.0	3010	2	O92979 hepatitis c
20	167	100.0	3010	2	O92979 hepatitis c
21	167	100.0	3010	2	O92979 hepatitis c
22	167	100.0	3010	2	O92979 hepatitis c
23	167	100.0	3010	2	O92979 hepatitis c
24	167	100.0	3010	2	O92979 hepatitis c
25	167	100.0	3010	2	O92979 hepatitis c
26	167	100.0	3010	2	O92979 hepatitis c
27	167	100.0	3010	2	O92979 hepatitis c
28	167	100.0	3010	2	O92979 hepatitis c
29	167	100.0	3010	2	O92979 hepatitis c
30	167	100.0	3010	2	O92979 hepatitis c
31	167	100.0	3010	2	O92979 hepatitis c

RESULT 1

ID	Accession	Length	Score	Query Match	Description
O41809	O41809	167	100.0	3010	PRELIMINARY; PRT; 1805 AA.
AC	O41809	167	100.0	3010	01-JAN-1998 (TrEMBLrel. 05, Created)
DT	01-JAN-1998	167	100.0	3010	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT	01-MAR-2004	167	100.0	3010	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Polyprotein.	167	100.0	3010	Hepatitis C virus.
OS	Hepatitis C virus.	167	100.0	3010	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;	167	100.0	3010	Hepatitis C virus.
OX	Hepatitis C virus.	167	100.0	3010	NCBI_TaxID=11103;
RN	[1]	167	100.0	3010	SEQUENCE FROM N.A.
RP	SEQUENCE FROM N.A.	167	100.0	3010	MEDLINE=97032798; PubMed=8878547; DOI=10.1006/bbrc.1996.1540;
RX	MEDLINE=97032798; PubMed=8878547; DOI=10.1006/bbrc.1996.1540;	167	100.0	3010	Yeh C.T., Chu C.M., Liaw Y.F.;
RA	Yeh C.T., Chu C.M., Liaw Y.F.;	167	100.0	3010	"Distinct composition of viral quasiespecies between ascites and serum
RT	"Distinct composition of viral quasiespecies between ascites and serum	167	100.0	3010	samples from patients with late stage chronic hepatitis C.";
RL	Biophys. Res. Commun. 227:524-529(1996).	167	100.0	3010	[2]
RP	SEQUENCE FROM N.A.	167	100.0	3010	MEDLINE=98033184; PubMed=9367361;
RX	MEDLINE=98033184; PubMed=9367361;	167	100.0	3010	Yeh C.T., Lu S.J., Lu S.J., Liaw Y.F.;
RA	Yeh C.T., Lu S.J., Lu S.J., Liaw Y.F.;	167	100.0	3010	"Molecular cloning of a defective hepatitis C virus genome from the
RT	"Molecular cloning of a defective hepatitis C virus genome from the	167	100.0	3010	ascitic fluid of a patient with hepatocellular carcinoma.";
RL	J. Gen. Virol. 78:0-0(0).	167	100.0	3010	[3]
RP	SEQUENCE FROM N.A.	167	100.0	3010	Yeh C.T., Chu C.M., Lu S.J., Liaw Y.F.;
RX	Yeh C.T., Chu C.M., Lu S.J., Liaw Y.F.;	167	100.0	3010	Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RA	Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.	167	100.0	3010	EMBL; U99019; AAB88251.1;
RL	EMBL; U99019; AAB88251.1;	167	100.0	3010	PIR; A61196; A61196.
RP	PIR; A61196; A61196.	167	100.0	3010	HSP; P26663; 10UV.
RX	HSP; P26663; 10UV.	167	100.0	3010	GO; GO:0005524; F:ATP binding; IEA.
RA	GO; GO:0005524; F:ATP binding; IEA.	167	100.0	3010	GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
RL	GO; GO:0008026; F:ATP-dependent helicase activity; IEA.	167	100.0	3010	GO; GO:0003723; F:RNA binding; IEA.
RP	GO; GO:0003723; F:RNA binding; IEA.	167	100.0	3010	GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
RX	GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.	167	100.0	3010	GO; GO:0006350; P:transcription; IEA.
RA	GO; GO:0006350; P:transcription; IEA.	167	100.0	3010	GO; GO:0019079; P:viral genome replication; IEA.
RL	GO; GO:0019079; P:viral genome replication; IEA.	167	100.0	3010	InterPro; IPR000345; CytC_heme_BS.
RP	InterPro; IPR000345; CytC_heme_BS.	167	100.0	3010	InterPro; IPR001410; DEAD.
RX	InterPro; IPR001410; DEAD.	167	100.0	3010	InterPro; IPR011545; DEAD/DEAH_N.
RA	InterPro; IPR011545; DEAD/DEAH_N.	167	100.0	3010	InterPro; IPR000745; HCV_NS4a.
RL	InterPro; IPR000745; HCV_NS4a.	167	100.0	3010	InterPro; IPR001490; HCV_NS4b.
RP	InterPro; IPR001490; HCV_NS4b.	167	100.0	3010	InterPro; IPR002868; HCV_NS5a.
RX	InterPro; IPR002868; HCV_NS5a.	167	100.0	3010	InterPro; IPR002166; HCV_RdRP.
RA	InterPro; IPR002166; HCV_RdRP.	167	100.0	3010	InterPro; IPR001650; Helicase_C.
RL	InterPro; IPR001650; Helicase_C.	167	100.0	3010	InterPro; IPR007095; RNA_pol_DS_PS.
RP	InterPro; IPR007095; RNA_pol_DS_PS.	167	100.0	3010	InterPro; IPR007094; RNA_pol_PSVir.
RX	InterPro; IPR007094; RNA_pol_PSVir.	167	100.0	3010	Pfam; PF01006; HCV_NS4a; 1.
RA	Pfam; PF01006; HCV_NS4a; 1.	167	100.0	3010	Pfam; PF01003; HCV_NS4b; 1.
RL	Pfam; PF01003; HCV_NS4b; 1.	167	100.0	3010	Pfam; PF01506; HCV_NS5a; 1.
RP	Pfam; PF01506; HCV_NS5a; 1.	167	100.0	3010	Pfam; PF00271; Helicase_C; 1.
RX	Pfam; PF00271; Helicase_C; 1.	167	100.0	3010	Pfam; PF00998; Viral_RdRP; 1.

```

DR SMART; SMO0487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Polypeptidein.
SQ SEQUENCE 1805 AA; 195957 MW; 28D2FD07D0F5074B CRC64;

Query Match 100.0%; Score 167; DB 2; Length 1805;
Best Local Similarity 100.0%; Pred. No. 3.6e-15;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNRLIASFASRGNHVSPHYVPESDAAARVTQIL 33
|||||
715 MNRLIASFASRGNHVSPHYVPESDAAARVTQIL 747

RESULT 2
Q7T4V8
ID Q7T4V8 PRELIMINARY; PRT; 1984 AA.
AC Q7T4V8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NS protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22694192; PubMed=12810084; DOI=10.1016/S0006-291X(03)01047-7;
RA Kato N., Sugiyama K., Namba K., Danesako H., Nakamura T., Takami M.,
RA Naka K., Nozaki A., Shimotohno K.;
RT "Establishment of a hepatitis C virus subgenomic replicon derived from
RT human hepatocytes infected in vitro.";
RL Biochem. Biophys. Res. Commun. 306:756-766(2003).
DR EMBL; AB109543; BAC7767.1; -.
DR HSSP; Q81755; LDXP.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RDRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR007095; RNA_pol_DS_Ps.
DR InterPro; IPR007094; RNA_pol_Psvir.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral_RDRP; 1.
DR SMART; SMO0487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
FT NON TER 1
SQ SEQUENCE 1984 AA; 214359 MW; 8AA8198D2C7B291C CRC64;

Query Match 100.0%; Score 167; DB 2; Length 1884;
Best Local Similarity 100.0%; Pred. No. 4e-15;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNRLIASFASRGNHVSPHYVPESDAAARVTQIL 33
|||||
715 MNRLIASFASRGNHVSPHYVPESDAAARVTQIL 747

```

```

Db 890 MNRLIASFASRGNHVSPHYVPESDAAARVTQIL 922
|||||
RESULT 3
Q81817
ID Q81817 PRELIMINARY; PRT; 2284 AA.
AC Q81817;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide precursor.
OS Hepatitis C virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=40271;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94068484; PubMed=7504283;
RA Hijikata M., Mizushima H., Tanji Y., Komada Y., Hirowatari Y.,
RA Akagi T., Kimura K., Shimotohno K.;
RT "Proteolytic processing and membrane association of putative
RT nonstructural proteins of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:10773-10777(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94333810; PubMed=8056334; DOI=10.1016/0378-1119(94)90008-6;
RA Tanji Y., Hijikata M., Hirowatari Y., Shimotohno K.;
RT "Identification of the domain required for trans-cleavage activity of
RT hepatitis C viral serine proteinase.";
RL Gene 145:215-219(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95056078; PubMed=7966638;
RA Tanji Y., Hijikata M., Hirowatari Y., Shimotohno K.;
RT "Hepatitis C virus polyprotein processing: kinetics and mutagenic
RT analysis of serine proteinase-dependent cleavage.";
RL J. Virol. 68:8418-8422(1994).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=95156583; PubMed=7853491;
RA Tanji Y., Hijikata M., Satoh S., Kaneko T., Shimotohno K.;
RT "Hepatitis C virus-encoded nonstructural protein NS4A has versatile
RT functions in viral protein processing.";
RL J. Virol. 69:1575-1581(1995).
DR EMBL; D16435; BAA03905.1; -.
DR PIR; A61196; A61196.
DR PIR; P0246; P0246.
DR PIR; PS0329; PS0329.
DR HSSP; Q81755; LDXP.
DR GO; GO:0019012; C:virion; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006508; P:transcription; IEA.
DR GO; GO:0006350; P:viral genome replication; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RDRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_Ps.
DR InterPro; IPR007094; RNA_pol_Psvir.

```

DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; Viral RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS0190; CYTOCHROME_C; UNKNOWN 1.
KW Nonstructural protein; Polyprotein; Signal.
FT SIGNAL 4 20 Potential.
FT CHAIN 55 83
FT CHAIN 986 1246 NS4B.
FT CHAIN 1247 1693 NS5A.
FT CHAIN 1694 2284 NS5B.
FT CHAIN 84 300 NS2.
FT CHAIN 21 54 Potential.
FT CHAIN 301 931 NS3.
FT CHAIN 932 985 NS4A.
SQ SEQUENCE 2284 AA; 247213 MW; DC272A1517046337 CRC64;

Query Match 100.0%; Score 167; DB 2; Length 2284;
Best Local Similarity 100.0%; Pred. No. 4.6e-15;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MNRLIAFASRGNHVSPTHYPESDAAARVTOIL 33
Db 1190 MNRLIAFASRGNHVSPTHYPESDAAARVTOIL 1222

RESULT 4
O92973 PRELIMINARY; PRT; 2864 AA.
ID O92973
AC O92973;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
infectious in vivo";
RL Virology 244:161-172(1998);
DR EMBL; AF054253; AAC15727.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q02828; INB4.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0003909; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR00345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV_capsid.

DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_Psvir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 2864 2864
SQ SEQUENCE 2864 AA; 310467 MW; 916DDA2FD0449C98 CRC64;
Query Match 100.0%; Score 167; DB 2; Length 2864;
Best Local Similarity 100.0%; Pred. No. 5.9e-15;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MNRLIAFASRGNHVSPTHYPESDAAARVTOIL 33
Db 1916 MNRLIAFASRGNHVSPTHYPESDAAARVTOIL 1948
RESULT 5
O92974 PRELIMINARY; PRT; 2864 AA.
ID O92974
AC O92974;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
infectious in vivo";
RL Virology 244:161-172(1998);
DR EMBL; AF054257; AAC15730.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q02828; INB4.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0003909; P:viral genome replication; IEA.
DR InterPro; IPR00345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV_capsid.

```

DR GO: 0006508; P:proteolysis and peptidolysis; IEA.
DR GO: 0006350; P:transcription; IEA.
DR GO: 0019079; P:viral genome replication; IEA.
DR GO: 0019087; P:viral transformation; IEA.
DR InterPro: IPR000345; CytC_heme_BS.
DR InterPro: IPR011545; DEAD/DEAH_N.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR InterPro: IPR000745; HCV NS4a.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV RdRp.
DR InterPro: IPR004109; Peptidase_S29.
DR InterPro: IPR009003; Pept_Ser_Cys.
DR InterPro: IPR002518; Pept_U39_HCV_NS2.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV NS1; 1.
DR Pfam: PF01538; HCV NS2; 1.
DR Pfam: PF02907; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00998; Viral RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON TER 2864 2864
SQ SEQUENCE 2864 AA; 310482 MW; 32CF23B5E59C4E59 CRC64;

Query Match 100.0%; Score 167; DB 2; Length 2864;
Best Local Similarity 100.0%; Pred. No. 5.9e-15;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNRLIAPASRGNHVSPHYVPESDAARVTQIL 33
Db 1916 MNRLIAPASRGNHVSPHYVPESDAARVTQIL 1948

RESULT 6
O92975 PRELIMINARY; PRT; 2864 AA.
AC O92975;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
infectious in vivo.";
RL Virology 244.161-172(1998).
DR EMBL; AF054258; AAC15731.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSSP; Q02828; 1NB4.
DR GO: 0016021; C:integral to membrane; IEA.

```

```

DR GO: 0019028; C:viral capsid; IEA.
DR GO: 0019031; C:viral envelope; IEA.
DR GO: 000524; F:ATP binding; IEA.
DR GO: 0008026; F:ATP-dependent helicase activity; IEA.
DR GO: 0003723; F:RNA binding; IEA.
DR GO: 0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO: 0008236; F:serine-type peptidase activity; IEA.
DR GO: 0005198; F:structural molecule activity; IEA.
DR GO: 0006508; P:proteolysis and peptidolysis; IEA.
DR GO: 0006350; P:transcription; IEA.
DR GO: 0019079; P:viral genome replication; IEA.
DR GO: 0019087; P:viral transformation; IEA.
DR InterPro: IPR000345; CytC_heme_BS.
DR InterPro: IPR011410; DEAD/DEAH_N.
DR InterPro: IPR011545; HCV capsid.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002519; HCV core.
DR InterPro: IPR002531; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR InterPro: IPR000745; HCV NS4a.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV RdRp.
DR InterPro: IPR004109; Peptidase_S29.
DR InterPro: IPR009003; Pept_Ser_Cys.
DR InterPro: IPR002518; Pept_U39_HCV_NS2.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV NS1; 1.
DR Pfam: PF01538; HCV NS2; 1.
DR Pfam: PF02907; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00998; Viral RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON TER 2864 2864
SQ SEQUENCE 2864 AA; 310416 MW; COCD3933ED07C6A5 CRC64;

Query Match 100.0%; Score 167; DB 2; Length 2864;
Best Local Similarity 100.0%; Pred. No. 5.9e-15;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNRLIAPASRGNHVSPHYVPESDAARVTQIL 33
Db 1916 MNRLIAPASRGNHVSPHYVPESDAARVTQIL 1948

RESULT 7
O92976 PRELIMINARY; PRT; 2864 AA.
AC O92976;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "transcripts of a chimeric clone of hepatitis C virus genotype 1b are

```


RT infectious in vivo.;
 RL Virology 244:161-172(1998)
 DR EMBL; AF054259; AAC15732.1; --
 DR PIR; A61196; A61196
 DR PIR; PQ0246; PQ0246
 DR PIR; PS0329; PS0329
 DR HSSP; Q02828; 1NB4
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0019028; C: viral capsid; IEA.
 DR GO; GO:0019031; C: viral envelope; IEA.
 DR GO; GO:0005524; F: ATP binding; IEA.
 DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F: RNA binding; IEA.
 DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
 DR GO; GO:0006508; F: proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; P: transcriptions; IEA.
 DR GO; GO:0019079; P: viral genome replication; IEA.
 DR GO; GO:0019087; P: viral transformation; IEA.
 DR InterPro; IPR000345; CytC_heme_BS.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR002522; HCV capsid.
 DR InterPro; IPR002521; HCV env.
 DR InterPro; IPR002519; HCV NS1.
 DR InterPro; IPR002531; HCV NS1.
 DR InterPro; IPR000745; HCV NS4a.
 DR InterPro; IPR001490; HCV NS4b.
 DR InterPro; IPR002868; HCV NS5a.
 DR InterPro; IPR002166; HCV RdRP.
 DR InterPro; IPR004109; Peptidase S29.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR InterPro; IPR002518; Pept_U39_HCV_NS2.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR Pfam; PF01543; HCV capsid; 1.
 DR Pfam; PF01542; HCV core; 1.
 DR Pfam; PF01539; HCV env; 1.
 DR Pfam; PF01560; HCV NS1; 1.
 DR Pfam; PF01538; HCV NS2; 1.
 DR Pfam; PF02907; HCV NS3; 1.
 DR Pfam; PF01006; HCV NS4a; 1.
 DR Pfam; PF01001; HCV NS4b; 1.
 DR Pfam; PF01506; HCV NS5a; 1.
 DR Pfam; PF00998; Viral_RdRP; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 FT NON_TER 2864 2864
 SQ SEQUENCE 2864 AA; 310481 MW; C6DEB415F4A1D1D6 CRC64;

Query Match 100.0%; Score 167; DB 2; Length 2864;
 Best Local Similarity 100.0%; Pred. No. 5.9e-15;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MNRLIAFASRGHNSPHTYVPESDAAARVTQIL 33
 |||||
 Db 1916 MNRLIAFASRGHNSPHTYVPESDAAARVTQIL 1948

RESULT 8
 Q9WLK8 PRELIMINARY; PRT; 2864 AA.
 ID Q9WLK8
 AC Q9WLK8;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-NAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.

OX NCBI_TaxID=11103;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=HC-J4;
 RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
 RA Yanagi M., St. Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
 RA Bukh J.;
 RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
 RT infectious in vivo."
 RL Virology 244:161-172(1998).
 DR EMBL; AF054255; AAC15729.1; --
 DR PIR; A61196; A61196
 DR PIR; PQ0246; PQ0246
 DR PIR; PS0329; PS0329
 DR HSSP; Q02828; 1NB4
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0019028; C: viral capsid; IEA.
 DR GO; GO:0019031; C: viral envelope; IEA.
 DR GO; GO:0005524; F: ATP binding; IEA.
 DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F: RNA binding; IEA.
 DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
 DR GO; GO:0006508; F: proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; P: transcriptions; IEA.
 DR GO; GO:0019079; P: viral genome replication; IEA.
 DR GO; GO:0019087; P: viral transformation; IEA.
 DR InterPro; IPR000345; CytC_heme_BS.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR002522; HCV capsid.
 DR InterPro; IPR002521; HCV env.
 DR InterPro; IPR002519; HCV NS1.
 DR InterPro; IPR002531; HCV NS1.
 DR InterPro; IPR000745; HCV NS4a.
 DR InterPro; IPR001490; HCV NS4b.
 DR InterPro; IPR002868; HCV NS5a.
 DR InterPro; IPR002166; HCV RdRP.
 DR InterPro; IPR004109; Peptidase S29.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR InterPro; IPR002518; Pept_U39_HCV_NS2.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR Pfam; PF01543; HCV capsid; 1.
 DR Pfam; PF01542; HCV core; 1.
 DR Pfam; PF01539; HCV env; 1.
 DR Pfam; PF01560; HCV NS1; 1.
 DR Pfam; PF01538; HCV NS2; 1.
 DR Pfam; PF02907; HCV NS3; 1.
 DR Pfam; PF01006; HCV NS4a; 1.
 DR Pfam; PF01001; HCV NS4b; 1.
 DR Pfam; PF01506; HCV NS5a; 1.
 DR Pfam; PF00998; Viral_RdRP; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 FT NON_TER 2864 2864
 SQ SEQUENCE 2864 AA; 310493 MW; C8AF2D0D7AE597E5 CRC64;

Query Match 100.0%; Score 167; DB 2; Length 2864;
 Best Local Similarity 100.0%; Pred. No. 5.9e-15;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MNRLIAFASRGHNSPHTYVPESDAAARVTQIL 33
 |||||
 Db 1916 MNRLIAFASRGHNSPHTYVPESDAAARVTQIL 1948

RESULT 9
 Q9WLK9 PRELIMINARY; PRT; 2864 AA.
 ID Q9WLK9

AC Q9WLK9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo.";
RL Virology 244:161-172(1998).
DR EMBL; AF054254; AAC15728.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSSP; Q02828; 1NB4.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; Viral RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 2864 2864
SQ SEQUENCE 2864 AA; 310550 MW; 20BC0F9F2940BEEA CRC64;

Query Match 100.0%; Score 167; DB 2; Length 2864;
Best Local Similarity 100.0%; Pred. No. 5.9e-15;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLIAPASRGNHVSPHYVPESDAARVTQIL 33
Db ||||| 1916 MNRLIAPASRGNHVSPHYVPESDAARVTQIL 1948

RESULT 10
Q9WLL0 PRELIMINARY; PRT; 2864 AA.
ID Q9WLL0;
AC Q9WLL0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo.";
RL Virology 244:161-172(1998).
DR EMBL; AF054252; AAC15726.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSSP; Q02828; 1NB4.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; Viral RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.

KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 FT NON_TER 2864 2864
 SQ SEQUENCE 2864 AA; 310557 MW; 53464C5C744191D9 CRC64;

Query Match 100.0%; Score 167; DB 2; Length 2864;
 Best Local Similarity 100.0%; Pred. No. 5.9e-15;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLIAPASRGNHVSPHYVPESDAAARVTQIL 33
 DB 1916 MNRLIAPASRGNHVSPHYVPESDAAARVTQIL 1948

RESULT 11
 Q9J3F4 PRELIMINARY; PRT; 3008 AA.
 ID Q9J3F4
 AC Q9J3F4
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polyprotein.
 GN Name=MD34;
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN SEQUENCE FROM N.A.
 RC STRAIN=MD34;
 RA Nagayama K., Kurosa M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
 RL Submitted (Nov-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF208024; AAF61205.1;
 DR PIR; A61196; A61196.
 DR PIR; PQ0246; PQ0246.
 DR PIR; PS0329; PS0329.
 DR HSP; O8JY51; 1CWX.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006350; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0019079; P:vital genome replication; IEA.
 DR GO; GO:0019087; P:viral transformation; IEA.
 DR InterPro; IPR000345; CytC_heme_BS.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR002522; HCV capsid.
 DR InterPro; IPR002521; HCV core.
 DR InterPro; IPR002519; HCV env.
 DR InterPro; IPR002531; HCV NS1.
 DR InterPro; IPR000745; HCV NS4b.
 DR InterPro; IPR001490; HCV NS4b.
 DR InterPro; IPR002868; HCV NS5a.
 DR InterPro; IPR002166; HCV RdRp.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR004109; Peptidase_S29.
 DR InterPro; IPR009003; Pept Ser Cys.
 DR InterPro; IPR002518; Pept U39 HCV NS2.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PVir.
 DR Pfam; PF01543; HCV_capsid; 1.
 DR Pfam; PF01542; HCV_core; 1.
 DR Pfam; PF01539; HCV env; 1.
 DR Pfam; PF01560; HCV NS1; 1.
 DR Pfam; PF01538; HCV NS2; 1.
 DR Pfam; PF02907; HCV_NS3; 1.

DR Pfam; PF01006; HCV_NS4a; 1.
 DR Pfam; PF01001; HCV_NS4b; 1.
 DR Pfam; PF01506; HCV_NS5a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00998; Viral_RdRp; 1.
 DR SMART; SMO0487; DEXOC; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 SQ SEQUENCE 3008 AA; 326835 MW; 99AE09E14C3109F4 CRC64;

Query Match 100.0%; Score 167; DB 2; Length 3008;
 Best Local Similarity 100.0%; Pred. No. 6.2e-15;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLIAPASRGNHVSPHYVPESDAAARVTQIL 33
 DB 1914 MNRLIAPASRGNHVSPHYVPESDAAARVTQIL 1946

RESULT 12
 POLG HCVBK
 ID POLG HCVBK STANDARD; PRT; 3010 AA.
 AC P26663;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
 DE Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
 DE (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
 DE (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
 DE NS4B (P27); Nonstructural protein NS4A (P4); Nonstructural protein
 DE NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
 OS Hepatitis C virus (isolate BK) (HCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11105;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=91140698; PubMed=1847440;
 RA Takamiwara A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J.,
 Onishi E., Andoh T., Yoshida I., Okayama H.;
 RT "Structure and organization of the hepatitis C virus genome isolated
 from human carriers." J. Virol. 65:1105-1113 (1991).
 RL J. Virol. 65:1105-1113 (1991).
 RN SEQUENCE OF 1487-1500.
 RP MEDLINE=96235224; PubMed=8647104;
 RA Borowski P., Heiland M., Oehlmann K., Becker B., Kornetky L.;
 RT "Non-structural protein 3 of hepatitis C virus inhibits
 phosphorylation mediated by cAMP-dependent protein kinase." Eur. J. Biochem. 237:611-618 (1996).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
 RX MEDLINE=97015088; PubMed=8861916; DOI=10.1016/S0092-8674(00)81350-1;
 RA Love R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
 Moenaw E.W., Adachi T., Hostomsky Z.;
 RT "The crystal structure of hepatitis C virus NS3 proteinase reveals a
 trypsin-like fold and a structural zinc binding site." Cell 87:331-342 (1996).
 RL Cell 87:331-342 (1996).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.
 RX MEDLINE=98227846; PubMed=9568891;
 RA Yan Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M.,
 Steinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.;
 RT "Complex of NS3 protease and NS4A peptide of BK strain hepatitis C
 virus: a 2.2-A resolution structure in a hexagonal crystal form." Protein Sci. 7:837-847 (1998).
 RL Protein Sci. 7:837-847 (1998).
 CC -!- FUNCTION: The small proteins NS2A, NS2B, NS4A and NS4B are
 hydrophobic, suggesting a possible membrane-related function. NS3
 and NS5 may play a role in the viral RNA replication.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral

CC precursor polypeptide, commonly with Asp or Glu in the P6
 CC position, Cys or Thr in P1 and Ser or Ala in P1'.
 CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA} (N).
 CC -|- SUBUNIT: The virion of this virus is a nucleocapsid covered by a
 CC lipoprotein envelope. The envelope consists of two proteins:
 CC protein M and glycoprotein E. The nucleocapsid is a complex of
 CC protein C and mRNA.
 CC -|- SIMILARITY: Contains 1 peptidase S29 domain.
 CC -|- SIMILARITY: Contains 1 peptidase U39 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M58335; AAA72945.1; -.
 CC PIR; A38465; GNVVTC.
 CC PDB; 1AIQ; X-ray; A/B/C=1027-1215.
 CC PDB; 1C2P; X-ray; A/B=2414-2989.
 CC PDB; 1CSJ; X-ray; A/B=2420-2950.
 CC PDB; 1CU1; X-ray; A/B=1013-1657.
 CC PDB; 1GX5; X-ray; A=2420-2955.
 CC PDB; 1GX6; X-ray; A=2420-2950.
 CC PDB; 1JXP; X-ray; A/B=1027-1212, C/D=1677-1692.
 CC PDB; 1NS3; X-ray; A/B=1027-1212.
 CC PDB; 1OUV; X-ray; A=2420-2997.
 CC PDB; 8OHM; X-ray; @=1216-1650.
 CC MEROPS; S29.001; -.
 CC InterPro; IPR001410; DEAD.
 CC InterPro; IPR002522; HCV capsid.
 CC InterPro; IPR002521; HCV core.
 CC InterPro; IPR002519; HCV env.
 CC InterPro; IPR002531; HCV NS1.
 CC InterPro; IPR000745; HCV NS4a.
 CC InterPro; IPR001490; HCV NS4b.
 CC InterPro; IPR002868; HCV NS5a.
 CC InterPro; IPR002166; HCV RdRp.
 CC InterPro; IPR009003; Pept_Ser_Cys.
 CC InterPro; IPR002518; Pept_U39_HCV_NS2.
 CC InterPro; IPR004109; Peptidase_S29.
 CC InterPro; IPR007095; RNA_pol_PS_P.
 CC InterPro; IPR007094; RNA_pol_PSVir.
 CC Pfam; PF01543; HCV capsid; 1.
 CC Pfam; PF01542; HCV_core; 1.
 CC Pfam; PF01539; HCV_env; 1.
 CC Pfam; PF01560; HCV_NS1; 1.
 CC Pfam; PF01538; HCV_NS2; 1.
 CC Pfam; PF02907; HCV_NS3; 1.
 CC Pfam; PF01006; HCV_NS4a; 1.
 CC Pfam; PF01001; HCV_NS4b; 1.
 CC Pfam; PF01506; HCV_NS5a; 1.
 CC Pfam; PF00998; Viral_RdRp; 1.
 CC SMART; SM00487; DEXDC; 1.
 CC 3D-structure; ATP-binding; Coat protein; Core protein;
 CC Direct protein sequencing; Envelope protein; Glycoprotein; Helicase;
 CC Hydrolase; Nonstructural protein; Polypeptide; Polypeptide;
 CC RNA-directed RNA polymerase; Serine protease; Transferase;
 CC Transmembrane.
 CC INIT_MET 1 1 Removed from capsid protein C by the
 CC cellular aminopeptidase.
 CC CHAIN 1 115 Capsid protein C (Potential).
 CC CHAIN 116 191 Matrix protein (Potential).
 CC CHAIN 192 383 Major envelope protein E (Potential).
 CC CHAIN 384 729 Nonstructural protein NS1/E2 (Potential).
 CC CHAIN 730 1006 Nonstructural protein NS2 (Potential).
 CC CHAIN 1007 1615 Protease/helicase NS3 (Potential).
 CC CHAIN 1616 1862 Nonstructural protein NS4 (Potential).
 CC CHAIN 1863 2013 Nonstructural protein NS4B (Potential).
 CC FT

FT CHAIN 2014 3010 RNA-directed RNA polymerase (Potential).
 FT TRANSMEM 347 369 Potential.
 FT ACT_SITE 1083 1083 Charge relay system.
 FT ACT_SITE 1107 1107 Charge relay system.
 FT ACT_SITE 1165 1165 Charge relay system.
 FT NP_BIND 1230 1237 ATP (Potential).
 FT SITE 1316 1319 DECH box.
 FT CARBOHYD 196 196 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 209 209 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 234 234 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 250 250 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 305 305 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 417 417 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 423 423 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 430 430 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 448 448 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 532 532 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 540 540 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 556 556 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 576 576 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 623 623 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 645 645 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2041 2041 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2077 2077 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2240 2240 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2529 2529 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2788 2788 N-linked (GlcNAc...) (Potential).
 FT STRAND 1031 1035 N-linked (GlcNAc...) (Potential).
 FT HELIX 1039 1047 N-linked (GlcNAc...) (Potential).
 FT STRAND 1050 1050 N-linked (GlcNAc...) (Potential).
 FT STRAND 1063 1063 N-linked (GlcNAc...) (Potential).
 FT STRAND 1068 1074 N-linked (GlcNAc...) (Potential).
 FT TURN 1075 1076 N-linked (GlcNAc...) (Potential).
 FT STRAND 1077 1081 N-linked (GlcNAc...) (Potential).
 FT HELIX 1082 1085 N-linked (GlcNAc...) (Potential).
 FT TURN 1086 1087 N-linked (GlcNAc...) (Potential).
 FT STRAND 1090 1092 N-linked (GlcNAc...) (Potential).
 FT TURN 1093 1094 N-linked (GlcNAc...) (Potential).
 FT STRAND 1095 1097 N-linked (GlcNAc...) (Potential).
 FT STRAND 1101 1103 N-linked (GlcNAc...) (Potential).
 FT TURN 1104 1107 N-linked (GlcNAc...) (Potential).
 FT STRAND 1108 1112 N-linked (GlcNAc...) (Potential).
 FT STRAND 1120 1120 N-linked (GlcNAc...) (Potential).
 FT STRAND 1122 1122 N-linked (GlcNAc...) (Potential).
 FT STRAND 1129 1133 N-linked (GlcNAc...) (Potential).
 FT TURN 1135 1136 N-linked (GlcNAc...) (Potential).
 FT STRAND 1139 1144 N-linked (GlcNAc...) (Potential).
 FT STRAND 1149 1157 N-linked (GlcNAc...) (Potential).
 FT HELIX 1158 1161 N-linked (GlcNAc...) (Potential).
 FT TURN 1162 1163 N-linked (GlcNAc...) (Potential).
 FT TURN 1165 1166 N-linked (GlcNAc...) (Potential).
 FT STRAND 1168 1171 N-linked (GlcNAc...) (Potential).
 FT TURN 1172 1174 N-linked (GlcNAc...) (Potential).
 FT STRAND 1175 1186 N-linked (GlcNAc...) (Potential).
 FT TURN 1187 1188 N-linked (GlcNAc...) (Potential).
 FT STRAND 1189 1197 N-linked (GlcNAc...) (Potential).
 FT HELIX 1198 1202 N-linked (GlcNAc...) (Potential).
 FT TURN 1203 1204 N-linked (GlcNAc...) (Potential).
 FT STRAND 1680 1688 N-linked (GlcNAc...) (Potential).
 FT STRAND 2421 2421 N-linked (GlcNAc...) (Potential).
 FT STRAND 2423 2425 N-linked (GlcNAc...) (Potential).
 FT STRAND 2439 2439 N-linked (GlcNAc...) (Potential).
 FT HELIX 2444 2449 N-linked (GlcNAc...) (Potential).
 FT STRAND 2453 2455 N-linked (GlcNAc...) (Potential).
 FT HELIX 2456 2458 N-linked (GlcNAc...) (Potential).
 FT STRAND 2461 2463 N-linked (GlcNAc...) (Potential).
 FT HELIX 2464 2471 N-linked (GlcNAc...) (Potential).
 FT STRAND 2474 2474 N-linked (GlcNAc...) (Potential).
 FT HELIX 2481 2494 N-linked (GlcNAc...) (Potential).
 FT TURN 2495 2496 N-linked (GlcNAc...) (Potential).
 FT STRAND 2498 2498 N-linked (GlcNAc...) (Potential).
 FT HELIX 2504 2509 N-linked (GlcNAc...) (Potential).
 FT TURN 2510 2510 N-linked (GlcNAc...) (Potential).

FT TURN 2513 2514
 FT TURN 2519 2520
 FT HELIX 2524 2528

Query Match 100.0%; Score 167; DB 1; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 6.3e-15;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLTAFASGRGNHVSPTHVPSDAAARVTQIL 33
 |||||
 DB 1916 MRLTAFASGRGNHVSPTHVPSDAAARVTQIL 1948

RESULT 13

POLG HCVA
 ID POLG HCVA STANDARD; PRT; 3010 AA.
 AC P26662;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
 DE Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
 DE (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
 DE (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
 DE NS4B (P27); Nonstructural protein NS4A (P4); Nonstructural protein
 DE NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)]
 OS Hepatitis C virus (isolate Japanese) (HCV)
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus
 OX NCBI_TaxID=11116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91088550; PubMed=2175903;
 RA Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
 RA Sugimura T., Shimotohno K.;
 RT "Molecular cloning of the human hepatitis C virus genome from Japanese
 RT patients with non-A, non-B hepatitis."
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528 (1990).
 RN [2]
 RP DISCUSSION OF SEQUENCE.
 RX MEDLINE=91192160; PubMed=1849488; DOI=10.1016/0014-5793(91)80322-T;
 RA Kato N., Hijikata M., Nakagawa M., Ootsuyama Y., Muraishi K.,
 RA Ohkoshi S., Shimotohno K.;
 RT "Molecular structure of the Japanese hepatitis C viral genome.";
 RL FEBS Lett. 280:325-328 (1991).
 CC -1- FUNCTION: The small proteins NS2A, NS2B, NS4A and NS4B are
 CC hydrophobic, suggesting a possible membrane-related function. NS3
 CC and NS5 may play a role in the viral RNA replication.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
 CC precursor polyprotein, commonly with Asp or Glu in the P6
 CC position, Cys or Thr in P1 and Ser or Ala in P1'.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA} (N).
 CC -1- SUBUNIT: The virion of this virus is a nucleocapsid covered by a
 CC lipoprotein envelope. The envelope consists of two proteins:
 CC protein M and glycoprotein E. The nucleocapsid is a complex of
 CC protein C and mRNA.
 CC -1- SIMILARITY: Contains 1 peptidase S29 domain.
 CC -1- SIMILARITY: Contains 1 peptidase U39 domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: D90208; BAA14233.1; -.
 CC PIR: A39253; GNWVCU.
 CC HSP: P26663; LJXP.
 CC MEROPS: S29.001; -.
 CC

DR MEROPS; U39.001; -.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR002522; HCV capsid.
 DR InterPro; IPR002521; HCV_core.
 DR InterPro; IPR002519; HCV_env.
 DR InterPro; IPR002531; HCV_NSI.
 DR InterPro; IPR000745; HCV_NS4a.
 DR InterPro; IPR001490; HCV_NS4b.
 DR InterPro; IPR002868; HCV_NS5a.
 DR InterPro; IPR002166; HCV_RdRP.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR InterPro; IPR002518; Pept_U39_Hcv_NS2.
 DR InterPro; IPR004109; Peptidase_S29.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR Pfam; PF01543; HCV_capsid; 1.
 DR Pfam; PF01542; HCV_core; 1.
 DR Pfam; PF01539; HCV_env; 1.
 DR Pfam; PF01560; HCV_NSI; 1.
 DR Pfam; PF01538; HCV_NS2; 1.
 DR Pfam; PF02907; HCV_NS3; 1.
 DR Pfam; PF01006; HCV_NS4a; 1.
 DR Pfam; PF01001; HCV_NS4b; 1.
 DR Pfam; PF01506; HCV_NS5a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00998; Viral_RdRP; 1.
 DR SMART; SM00487; DEXDC; 1.
 KW ATP-binding; Coat protein; Core protein; Envelope protein;
 KW Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
 KW RNA-directed RNA polymerase; Serine protease; Transferase;
 KW Transmembrane.
 FT INIT_MET 1 1 Removed from capsid protein C by the
 FT CHAIN 1 115 cellular aminopeptidase.
 FT CHAIN 116 191 Capsid protein C (Potential).
 FT CHAIN 192 383 Matrix protein (Potential).
 FT CHAIN 384 729 Major envelope protein E (Potential).
 FT CHAIN 730 1006 Nonstructural protein NS1 (Potential).
 FT CHAIN 1007 1615 Nonstructural protein NS2 (Potential).
 FT CHAIN 1616 1862 Protease/helicase NS3 (Potential).
 FT CHAIN 1863 2013 Nonstructural protein NS4A (Potential).
 FT CHAIN 2014 3010 Nonstructural protein NS4B (Potential).
 FT TRANSMEM 347 369 RNA-directed RNA polymerase (Potential).
 FT ACT_SITE 1083 1083 Charge relay system (By similarity).
 FT ACT_SITE 1107 1107 Charge relay system (By similarity).
 FT ACT_SITE 1165 1165 Charge relay system (By similarity).
 FT NP_BIND 1230 1237 ATP (Potential).
 FT SITE 1316 1319 DECH box.
 FT CARBOHYD 196 196 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 209 209 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 234 234 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 250 250 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 305 305 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 417 417 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 423 423 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 430 430 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 448 448 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 532 532 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 556 556 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 576 576 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 623 623 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 645 645 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2041 2041 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2077 2077 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2240 2240 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2788 2788 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 3010 AA; 327017 MW; AA993794F46DB185 CRC64;

Query Match 100.0%; Score 167; DB 1; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 6.3e-15;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLIAFASRGHNVSPHYVPESDAAARVTQIL 33
 DB 1916 MNRLIAFASRGHNVSPHYVPESDAAARVTQIL 1948

RESULT 14
 POLG_HCVJT
 ID POLG_HCVJT STANDARD; PRT; 3010 AA.
 AC Q00269;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
 DE Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
 DE (GP68) (NS1); Protein P7; Nonstructural protein NS2 (P21)
 DE (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
 DE (EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
 DE NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
 DE NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
 OS Hepatitis C virus (isolate HC-JT) (HCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=31642;
 RN [1]_TaxID=31642;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92295714; PubMed=1318627; DOI=10.1016/0168-1702(92)90066-I;
 RA Tanaka T., Kato N., Nakagawa M., Ooteyama Y., Cho M.U., Nakazawa T.,
 RA Hijikata M., Ishimura Y., Shimotohno K.;
 RT "Molecular cloning of hepatitis C virus genome from a single Japanese
 RT carrier: sequence variation within the same individual and among
 RT infected individuals.";
 RL Virus Res. 23:39-53(1992).
 CC -!- FUNCTION: The small proteins NS2A, NS2B, NS4A and NS4B are
 CC hydrophobic, suggesting a possible membrane-related function. NS3
 CC and NS5 may play a role in the viral RNA replication.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
 CC precursor polyprotein, commonly with Asp or Glu in the P6
 CC position, Cys or Thr in P1 and Ser or Ala in P1'.
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA} (N).
 CC -!- SUBUNIT: The virion of this virus is a nucleocapsid covered by a
 CC lipoprotein envelope. The envelope consists of two proteins:
 CC protein M and glycoprotein E. The nucleocapsid is a complex of
 CC protein C and mRNA.
 CC -!- SIMILARITY: Contains 1 peptidase S29 domain.
 CC -!- SIMILARITY: Contains 1 peptidase U39 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D11168; BAA01943.1; -;
 DR PIR; A45573; A45573.
 DR HSSP; P26663; 1JXP.
 DR MEROPS; S29.001; -;
 DR MEROPS; U39.001; -;
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR002522; HCV_capsid.
 DR InterPro; IPR002521; HCV_core.
 DR InterPro; IPR002519; HCV_env.
 DR InterPro; IPR002531; HCV_NS1.
 DR InterPro; IPR000745; HCV_NS4a.
 DR InterPro; IPR001490; HCV_NS4b.
 DR InterPro; IPR002868; HCV_NS5a.
 DR InterPro; IPR002166; HCV_NS5a.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR009003; Pept Ser Cys.
 DR InterPro; IPR002518; Pept U39_HCV_NS2.
 DR InterPro; IPR004109; Peptidase_S29.

DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR Pfam; PF01543; HCV_capsid; 1.
 DR Pfam; PF01545; HCV_core; 1.
 DR Pfam; PF01539; HCV_env; 1.
 DR Pfam; PF01560; HCV_NS1; 1.
 DR Pfam; PF01538; HCV_NS2; 1.
 DR Pfam; PF02907; HCV_NS3; 1.
 DR Pfam; PF01006; HCV_NS4a; 1.
 DR Pfam; PF01001; HCV_NS4b; 1.
 DR Pfam; PF01506; HCV_NS5a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00998; Viral_RdRP; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR ATP-binding; Coat protein; Core protein; Envelope protein; Polyprotein;
 KW Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Transferase;
 KW RNA-directed RNA polymerase; Serine protease; Transmembrane.
 FT INIT_MET 1 1 Removed from capsid protein C by the
 FT CHAIN 1 115 cellular aminopeptidase.
 FT CHAIN 116 191 Capsid protein C (Potential).
 FT CHAIN 192 383 Matrix protein (Potential).
 FT CHAIN 384 729 Major envelope protein E (Potential).
 FT CHAIN 730 1006 Nonstructural protein NS1/E2 (Potential).
 FT CHAIN 1007 1615 Nonstructural protein NS2 (Potential).
 FT CHAIN 1616 1862 Protease/helicase NS3 (Potential).
 FT CHAIN 1863 2013 Nonstructural protein NS4A (Potential).
 FT CHAIN 2014 3010 Nonstructural protein NS4B (Potential).
 FT CHAIN 3010 369 RNA-directed RNA polymerase (Potential).
 FT ACT_SITE 1083 1083 Charge relay system (By similarity).
 FT ACT_SITE 1107 1107 Charge relay system (By similarity).
 FT ACT_SITE 1165 1165 Charge relay system (By similarity).
 FT NP_BIND 1230 1337 App (Potential).
 FT SITE 1316 1319 DECH box.
 FT CARBOHYD 196 196 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 209 209 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 234 234 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 250 250 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 305 305 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 417 417 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 423 423 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 430 430 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 448 448 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 532 532 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 540 540 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 556 556 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 576 576 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 623 623 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 645 645 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2041 2041 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2077 2077 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2240 2240 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2529 2529 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 2788 2788 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 3010 AA; 326573 MW; 94A1C77435D642BB CRC64;
 Query Match 100.0%; Score 167; DB 1; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 6.3e-15;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLIAFASRGHNVSPHYVPESDAAARVTQIL 33
 DB 1916 MNRLIAFASRGHNVSPHYVPESDAAARVTQIL 1948

RESULT 15
 POLG_HCVTW
 ID POLG_HCVTW STANDARD; PRT; 3010 AA.
 AC P29846;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)

FT STRAND 1189 1197
FT HELIX 1198 1205
FT STRAND 1680 1688
SQ SEQUENCE 3010 AA; 327047 MW; AAD267D55CDFE215 CRC64;
Query Match 100.0%; Score 167; DB 1; Length 3010;
Best Local Similarity 100.0%; Pred. No. 6.3e-15; Indels 0; Gaps 0;
Matches 33; Conservative 0; Mismatches 0;
QY 1 MNRLIAFASRGNHVSPTHYVPESDAAARVTQIL 33
DB 1916 MNRLIAFASRGNHVSPTHYVPESDAAARVTQIL 1948

Search completed: August 12, 2005, 14:21:21
Job time : 31.4865 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:00:18 ; Search time 107.012 Seconds
(without alignments)
368.645 Million cell updates/sec

Title: US-09-758-308-5
Perfect score: 554
Sequence: 1 KATCTTHDSPADLIEANL.....NPPLLESWKDPDYVPVPHG 102

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	554	100.0	102	AAB31699	Antigenic
2	554	100.0	3010	AAR68622	HCV prote
3	554	100.0	3010	AAR68864	Hepatitis
4	554	100.0	3010	AAR82694	Partial H
5	554	100.0	3010	ADF88597	Hepatitis
6	521	94.0	392	AAR62868	Korean-ty
7	521	94.0	3010	AAR30616	Polypepti
8	520	93.9	299	AAR25118	Non-A, No
9	518	93.5	228	AAR10743	Non-A non
10	516	93.1	3010	AAR33417	Blood tra
11	514	92.8	1736	AAB36932	Hepatitis
12	514	92.8	1985	ABU09574	HCV Met-N
13	514	92.8	1985	ABU09575	HCV Met-N
14	514	92.8	1985	ADR38450	Hepatitis
15	514	92.8	2201	AARW01680	Hepatitis
16	514	92.8	3010	AARW01680	HCV NS2-N
17	514	92.8	3010	AAR20091	Non-A, no
18	514	92.8	3010	AAV06423	Non-A, no
19	512	92.4	3010	AAR34580	Human hep
20	511	92.2	447	ABB05639	Hepatitis
21	511	92.2	449	AAW37129	Hepatitis
22	511	92.2	483	AAR58590	Hepatitis
23	511	92.2	483	ADL17781	Hepatitis
24	511	92.2	3011	AAR34468	Encoded b
25	510	92.1	1038	AAE02357	Hepacivir

26	510	92.1	1038	4	AAE02356	Hepacivir
27	509	91.9	199	2	AAR10747	Non-A non
28	509	91.9	380	2	AAR10748	Non-A non
29	507	91.5	447	5	AAE15724	Hepatitis
30	507	91.5	447	5	AAE15725	Hepatitis
31	507	91.5	447	5	AAE15718	Hepatitis
32	507	91.5	447	5	AAE15723	Hepatitis
33	507	91.5	447	5	AAE15726	Hepatitis
34	507	91.5	447	5	AAE15721	Hepatitis
35	507	91.5	1985	5	AAO18001	Hepatitis
36	507	91.5	1985	5	AAE15729	Hepatitis
37	507	91.5	1985	5	AAE15731	Hepatitis
38	507	91.5	1985	5	AAE15720	Hepatitis
39	507	91.5	1985	5	AAE15717	Hepatitis
40	507	91.5	1985	5	AAE15727	Hepatitis
41	507	91.5	1985	5	AAE15728	Hepatitis
42	507	91.5	1985	5	AAE15722	Hepatitis
43	507	91.5	1985	5	AAE15730	Hepatitis
44	507	91.5	1985	8	ADJ57846	HCV repli
45	507	91.5	2063	7	ADD67963	Hepatitis

ALIGNMENTS

RESULT 1
AAB31699
ID AAB31699 standard; peptide; 102 AA.
XX
AC AAB31699;
XX
DT 30-APR-2001 (first entry)
XX
DE Antigenic epitope of the Hepatitis C virus (HCV) NS5a protein.
XX
KW Antigen; HCV; polyprotein; core protein; NS3 protein; NS4 protein;
NS4a protein; HCV infection.
XX
OS Hepatitis C virus.
XX
PN WO200104149-A1.
XX
PD 18-JAN-2001.
XX
PF 07-JUL-2000; 2000WO-US018704.
XX
PR 09-JUL-1999; 99WO-US015578.
XX
(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
DR WPI; 2001-138316/14.
XX
PT New (mosaic) polypeptides, useful as reagents in assays for the diagnosis
of monitoring of HCV infections and as components of anti-HCV vaccines,
comprises antigenic groups of hepatitis C virus (HCV).
XX
PS Claim 6; Page 39; 52pp; English.
XX
CC The present sequence represents an antigenic epitope from a hepatitis C
virus (HCV) NS5a protein. The peptide comprises amino acids 2322-2423 of
the HCV polyprotein. The specification describes antigenic epitopes from
the HCV core protein, NS3 protein, NS4 protein and NS4a protein. The
antigenic peptides are useful as diagnostic reagents for detecting HCV in
a biological sample. They are also useful for monitoring HCV infection in
a patient sample in addition to diagnosis. Pharmaceutical compositions
comprising the peptides are useful for preventing, minimizing or reducing
HCV infection in patients who have been exposed to HCV or to individuals,
such as health care workers or blood product recipients, who are more
likely to become exposed to HCV infection. The antigenic peptides are
also useful for generating antibodies which can be used to detect HCV
proteins in a sample or for laboratory research purposes

```
XX SQ Sequence 102 AA;
Query Match 100.0%; Score 554; DB 4; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.6e-56;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATCTTHDSDPADLIEANLLWRQEMGNITRVESENKVVILDSFDPPIRAVEDEREISVP 60
|
Db 1 KATCTTHDSDPADLIEANLLWRQEMGNITRVESENKVVILDSFDPPIRAVEDEREISVP 60
|

QY 61 AEILRKPRKFPALPIWARPDPYNPPLLESWKDPDYVPPVPHG 102
|
Db 61 AEILRKPRKFPALPIWARPDPYNPPLLESWKDPDYVPPVPHG 102
|

RESULT 2
AAR68622
ID AAR68622 standard; protein; 3010 AA.
AC AAR68622;
DT 16-OCT-2003 (revised)
DT 16-OCT-1995 (first entry)
XX HCV protein cleavable with new serine proteinase.
XX proteinase; serine; cleavage; hepatitis C virus; HCV.
XX Hepatitis C virus; Virus.
XX Key Location/Qualifiers
FH Key 2419..2420
FT Cleavage-site /note= "Serine protease cleavage site"
XX JP06315377-A.
XX 15-NOV-1994.
XX 06-MAY-1993; 93JP-00105666.
XX 06-MAY-1993; 93JP-00105666.
XX (KAEN/) KAENNO K.
PA (SUMQ ) SUMITOMO METAL IND LTD.
PA (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.
XX WPI; 1995-032330/05.
DR N-PSDB; AAQ80498.
XX
XX New HCV-originated proteinase active substance - used for site-specific
PT cleavage by an intermolecular reaction and the purification thereof.
XX Disclosure; Page 10-19; 23pp; Japanese.
XX This protein from HCV (hepatitis C virus) (encoded by AAQ80498) is
CC cleaved between amino acids 2419 and 2420, by a new serine protease,
CC contg. the sequence of AAR68621. The proteinase is purified as a fused
CC product with the dihydrofolate reductase protein by using a methotrexate
CC column. It can be used for the development of an inhibitor for HCV
CC proteinase. (Updated on 16-OCT-2003 to standardise OS field)
XX
XX SQ Sequence 3010 AA;
Query Match 100.0%; Score 554; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 1.1e-54;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATCTTHDSDPADLIEANLLWRQEMGNITRVESENKVVILDSFDPPIRAVEDEREISVP 60
|
Db 2212 KATCTTHDSDPADLIEANLLWRQEMGNITRVESENKVVILDSFDPPIRAVEDEREISVP 2271
|
QY 61 AEILRKPRKFPALPIWARPDPYNPPLLESWKDPDYVPPVPHG 102
|

Db 2272 AEILRKPRKFPALPIWARPDPYNPPLLESWKDPDYVPPVPHG 2313
|

RESULT 3
AAR68864
ID AAR68864 standard; protein; 3010 AA.
XX AAR68864;
XX 06-DEC-1995 (first entry)
XX Hepatitis C virus RNA helicase.
XX Hepatitis C virus HCV; non-A non-B; helicase gene; RNA helicase;
XX Hepatitis C virus; recombinant production.
XX Hepatitis C virus.
XX Key Location/Qualifiers
FH Key 196..198
FT /label= N-linked glycosylation site
FT 209..211
FT /label= N-linked glycosylation site
FT 234..236
FT /label= N-linked glycosylation site
FT 250..252
FT /label= N-linked glycosylation site
FT 305..307
FT /label= N-linked glycosylation site
FT 325..327
FT /label= N-linked glycosylation site
FT 417..419
FT /label= N-linked glycosylation site
FT 423..425
FT /label= N-linked glycosylation site
FT 430..432
FT /label= N-linked glycosylation site
FT 448..450
FT /label= N-linked glycosylation site
FT 532..534
FT /label= N-linked glycosylation site
FT 556..558
FT /label= N-linked glycosylation site
FT 576..578
FT /label= N-linked glycosylation site
FT 623..625
FT /label= N-linked glycosylation site
FT 645..647
FT /label= N-linked glycosylation site
FT 1213..1215
FT /label= N-linked glycosylation site
FT 1255..1257
FT /label= N-linked glycosylation site
FT 2041..2043
FT /label= N-linked glycosylation site
FT 2077..2079
FT /label= N-linked glycosylation site
FT 2240..2242
FT /label= N-linked glycosylation site
FT 2788..2790
FT /label= N-linked glycosylation site
XX JP06319583-A.
XX 22-NOV-1994.
XX 18-SEP-1992; 92JP-00249241.
XX 18-SEP-1992; 92JP-00249241.
XX (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.
```

DR WPI; 1995-040330/06.
 XX N-PSDB; AAQ81559.
 XX of hepatitis C virus helicase gene in baculovirus - useful for large
 PT scale prodn. of RNA helicase.
 XX
 XX Claim 1; Fig 1-4; 9pp; Japanese.
 CC AAQ81559 encodes AAR8864 hepatitis C virus (HCV) RNA helicase. The DNA
 CC was used in the construction of an expression vector, which was used to
 CC transform a baculovirus host. The transformed baculovirus could then be
 CC used for the recombinant prodn. of HCV RNA helicase
 XX
 XX Sequence 3010 AA;
 SQ
 Query Match 100.0%; Score 554; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 1.1e-54;
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDEIREISVP 60
 DB 2212 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDEIREISVP 2271
 QY 61 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVVG 102
 DB 2272 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVVG 2313
 RESULT 4
 AAR82694
 ID AAR82694 standard; protein; 3010 AA.
 XX
 AC AAR82694;
 XX
 DT 16-OCT-2003 (revised)
 DT 14-NOV-1996 (first entry)
 XX
 DE Partial HCV non-structural polyprotein.
 XX
 KW proteinase; hepatitis C virus; screening; inhibitor; proteolytic;
 KW identification; cleavage.
 XX
 OS Hepatitis C virus; Virus.
 XX
 FH Key Location/Qualifiers
 FT Protein 898..1233
 FT /note= "partial proteinase; see AAR82692"
 FT Protein 992..1907
 FT /note= "partial proteinase; see AAR82693"
 XX
 PN JP07184548-A
 XX
 PD 25-JUL-1995.
 XX
 XX 05-FEB-1993; 93JP-00018854.
 XX
 PR 07-FEB-1992; 92JP-00022657.
 PR 18-SEP-1992; 92JP-00249240.
 PR 04-DEC-1992; 92JP-00325303.
 XX
 XX (KAEN/) KAENNO K.
 PA (SUMO) SUMITOMO METAL IND LTD.
 PA (SOYA-) SOYAKU GIUTSU KENKYUSHO KK.
 XX
 WPI; 1995-287962/38.
 DR N-PSDB; AAT03960.
 XX
 XX An HCV proteinase active substance - which has activity as an anti-HCV
 PT agent and can be used to screen for proteinase inhibitors.
 XX
 PS Disclosure; Page 39-48; 52pp; Japanese.
 XX
 XX The present sequence is a partial Hepatitis C Virus (HCV) polyprotein

CC from the non-structural region. Partial proteinase sequences (AAR82692-93) are contained within this sequence. The proteinases can be used as anti-HCV agents. They can also be used to screen cpds. for their ability to inhibit their proteolytic activity. In this way proteinase inhibitors can be identified. (Updated on 16-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 3010 AA;
 Query Match 100.0%; Score 554; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 1.1e-54;
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDEIREISVP 60
 DB 2212 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDEIREISVP 2271
 QY 61 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVVG 102
 DB 2272 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVVG 2313
 RESULT 5
 ADF88597
 ID ADF88597 standard; protein; 3010 AA.
 XX
 AC ADF88597;
 XX
 DT 26-FEB-2004 (first entry)
 XX
 DE Hepatitis C virus NS3 gene protein, SEQ ID No 6.
 XX
 KW liver cancer; hepatitis-C virus; NS3 gene; carcinogenicity.
 XX
 OS Hepatitis C virus.
 XX
 PN JP2003210181-A.
 XX
 PD 29-JUL-2003.
 XX
 PF 30-MAY-2002; 2002JP-00158335.
 XX
 PR 16-NOV-2001; 2001JP-00352443.
 XX
 PA (SHIN-) ZH SHINSANGYO SOZO KENKYU KIKO.
 XX
 DR WPI; 2003-819836/77.
 DR N-PSDB; ADF88596.
 XX
 PT Diagnosing liver cancer, involves amplifying amino terminal region of
 PT hepatitis-C virus gene using predetermined primer and determining
 PT hepatitis-C virus in base sequence of obtained DNA fragment.
 XX
 PS Disclosure; SEQ ID NO 6; 36pp; Japanese.
 XX
 CC The invention relates to the novel testing method for diagnosing liver
 CC cancer. The novel method comprises amplifying the amino terminal region
 CC of a hepatitis-C virus NS3 gene using a predetermined primer and
 CC determining the hepatitis-C virus in a base sequence of the obtained DNA
 CC fragment. The novel testing method is useful for diagnosing liver cancer
 CC and also used in a gene amplification technique, a clinical laboratory
 CC test reagent, a polymerase chain reaction, a base sequence analysis and
 CC genetic engineering. The method enables the detection of a hepatitis-C
 CC virus having high carcinogenicity with high specificity. This sequence
 CC represents the protein of the hepatitis-C virus NS3 gene of the
 CC invention.
 XX
 SQ Sequence 3010 AA;
 Query Match 100.0%; Score 554; DB 7; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 1.1e-54;
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDEIREISVP 60

```

Db      2212 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESKVKVILDSFDPPIRAVEDEREISVP 2271
      61 ABILRKPRKFPFPALPIWAPDYNPPLESWKDPDYVPPVVG 102
      2272 ABILRKPRKFPFPALPIWAPDYNPPLESWKDPDYVPPVVG 2313

RESULT 6
AAR62868
ID AAR62868 standard; protein; 392 AA.
XX
AC AAR62868;
XX
DT 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 14-JUL-1995 (first entry)
XX
DE Korean-type Hepatitis C Virus non-structural protein 5-1.2 epitope.
XX
KW Korean-type hepatitis C virus; KHCV; non-A, non-B hepatitis; NANBH;
KW viral non-structural protein; NS5-1.2; epitope; antigen; immunodiagnosis.
XX
OS Hepatitis C virus; Virus (Korean-type).
XX
PN W09425874-A1.
XX
PD 10-NOV-1994.
XX
PF 27-APR-1994; 94WO-KR000039.
XX
PR 28-APR-1993; 93KR-00007231.
XX
PA (LUCK-) LUCKY CO LTD.
XX
PI Cho JM, Choi DY, Kim CH, So HS, Yang JY, Kim IS, Choi DS;
XX
DR WPI; 1994-358478/44.
DR N-PSDB; AAQ73204.
XX
PT Kit for simultaneous diagnosis of hepatitis B and C - comprising one or
PT more hepatitis B and C virus antigenic proteins including one or more
PT epitope(s).
XX
PS Claim 2; Fig 5; 89pp; English.
XX
CC The KHCV NS5-1.2 epitope is a preferred antigen for use in an
CC immunodiagnostic kit for simultaneous detection of hepatitis C and B
CC viruses. The kit comprises antigenic proteins from both viruses. The HCV
CC protein is pref. one of KHCV CORE 14, KHCV 897, KHCV NS4E, KHCV NS4B1E2
CC or KHCV NS5-1.2 proteins; the HBV protein is pref. HBV CORE or Pre S2 SAG
CC protein. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-OCT
CC -2003 to standardise OS field)
XX
SQ Sequence 392 AA;
      Query Match 94.0%; Score 521; DB 2; Length 392;
      Best Local Similarity 94.1%; Pred. No. 6e-52;
      Matches 96; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESKVKVILDSFDPPIRAVEDEREISVP 60
Db 110 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESKVKVILDSFDPPIRAVEDEREISVP 169
      61 AEILRKPRKFPFPALPIWAPDYNPPLESWKDPDYVPPVVG 102
      170 AEILRKPRKFPFPALPIWAPDYNPPLESWKDPDYVPPVVG 211

RESULT 7
AAR30616
ID AAR30616 standard; protein; 3010 AA.
XX

```

```

AC AAR30616;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 19-MAY-1993 (first entry)
XX
DE Polypeptide coded by Korean HCV full cDNA sequence LBC1.
XX
KW KHCV-LBC1; diagnosis; vaccine.
XX
OS Hepatitis C virus.
XX
PN EP521318-A2.
XX
PD 07-JAN-1993.
XX
PF 10-JUN-1992; 92EP-00109753.
XX
PR 10-JUN-1991; 91KR-00009510.
PR 06-AUG-1991; 91KR-00013601.
XX
PA (LUCK-) LUCKY LTD.
XX
PI Cho JM, Lee YB, Park YW, Lim KJ, Choi DY, So HS, Kim CH;
PI Kim ST, Yang JY;
XX
DR WPI; 1993-001883/01.
DR N-PSDB; AAQ33282.
XX
PT DNA and polypeptide(s) from a new type of hepatitis C virus (KHCV) - for
PT diagnosing and vaccinating against KHCV infections.
XX
PS Disclosure; Fig 2; 119pp; English.
XX
CC The polypeptide is that encoded by the full cDNA sequence of Korean
CC hepatitis C virus (KHCV) cDNA, KHCV-LBC1. It or its fragments may be used
CC in a specific and accurate method for detecting KHCV antibodies in the
CC serum of hepatitis C patients. Antibodies directed against these
CC polypeptides are useful for the purification of KHCV antigens and for the
CC development of an improved diagnostic to detect KHCV antigens in a
CC sample. The polypeptides may also be used in a vaccine for treatment and
CC prevention of KHCV infection at a dosage of 5-200 ug/peptide. (Updated on
CC 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
CC field.)
XX
SQ Sequence 3010 AA;
      Query Match 94.0%; Score 521; DB 2; Length 3010;
      Best Local Similarity 94.1%; Pred. No. 7.5e-51;
      Matches 96; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESKVKVILDSFDPPIRAVEDEREISVP 60
Db 2212 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESKVKVILDSFDPPIRAVEDEREISVP 2271
      61 ABILRKPRKFPFPALPIWAPDYNPPLESWKDPDYVPPVVG 102
      2272 ABILRKPRKFPFPALPIWAPDYNPPLESWKDPDYVPPVVG 2313

RESULT 8
AAR25118
ID AAR25118 standard; protein; 299 AA.
XX
AC AAR25118;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 07-DEC-1992 (first entry)
XX
DE Non-A, Non-B Hepatitis Virus antigen #9.
XX
KW Antigen S29; NANBH; Hepatitis C; HCV; T064; T069; T06A; ELISA.

```


FT FT Misc-difference 311 /label= Phe, Ser
FT FT /label= Gly, Asp
FT FT Misc-difference 398
FT FT /label= Ser, Arg, Gly
FT FT Misc-difference 400
FT FT /label= Thr, Ala
FT FT Misc-difference 405
FT FT /label= Gln, Pro, Leu
FT FT Misc-difference 410
FT FT /label= Lys, Arg
FT FT Misc-difference 418
FT FT /label= Gly, Asp
FT FT Misc-difference 430
FT FT /label= Asn, Asp
FT FT Misc-difference 438
FT FT /label= Phe, Leu
FT FT Misc-difference 478
FT FT /label= Arg, Lys
FT FT Misc-difference 759
FT FT /label= Leu, Val
FT FT Misc-difference 1017
FT FT /label= Ser, Asn
FT FT Misc-difference 1036
FT FT /label= Thr, Ala
FT FT Misc-difference 1056
FT FT /label= Glu, Asp
FT FT Misc-difference 1201
FT FT /label= Met, Thr
FT FT Misc-difference 1205
FT FT /label= Met, Ile
FT FT Misc-difference 1255
FT FT /label= Asn, Tyr
FT FT Misc-difference 1263
FT FT /label= Gly, Asp
FT FT Misc-difference 1455
FT FT /label= Asn, Asp
FT FT Misc-difference 1828
FT FT /label= Ala, Thr
FT FT Misc-difference 1895
FT FT /label= Gly, Arg
FT FT Misc-difference 1896
FT FT /label= Gly, Ile
FT FT Misc-difference 2143
FT FT /label= Glu, Val
FT FT Misc-difference 2144
FT FT /label= Asp, Glu
FT FT Misc-difference 2462
FT FT /label= Cys, Arg
FT FT Misc-difference 2486
FT FT /label= Val, Met
FT FT Misc-difference 2488
FT FT /label= Lys, Gln
FT FT Misc-difference 2844
FT FT /label= Leu, Met
FT FT Misc-difference 2862
FT FT /label= Leu, Gln
FT FT Misc-difference 2917
FT FT /label= Arg, Leu
FT FT Misc-difference 2968
FT FT /label= Ser, Gly
FT FT Misc-difference 2989
FT FT /label= Cys, Arg
FT FT Misc-difference 2990
FT FT /label= Tyr, Cys
XX
PN JP06105690-A.
XX
PD 19-APR-1994.
XX
XX 10-MAR-1992; 92JP-00051885.
XX
PR 10-MAR-1992; 92JP-00051885.

XX (KAEN/) KAENNO K.
PA
XX WPI; 1994-163130/20.
XX N-PSDB; AAQ63499.
XX
PT Blood-transmissible non-A non-B hepatitis virus DNA - used for detection
PT of hepatitis virus.
XX
XX Claim 1; Page 8-20; 22pp; Japanese.
XX
XX This sequence is encoded by the genome of a blood transmissible non-A,
CC non-B hepatitis (NANBH) virus. The cDNA sequence was isolated using the
CC primers given in AAQ63500-35. The amplified fragments are used in the
CC detection of hepatitis virus. The target DNA was isolated from serum of
CC chronically infected NANBH patients who were C100 antibody-positive and
CC HCV RNA (NS5 region) positive. Reverse transcription-PCR and PCR were
CC performed on cDNA and the total human NANBH DNA was constructed from 23
CC clones
XX
XX Sequence 3010 AA;
SQ
Query Match 93.1%; Score 516; DB 2; Length 3010;
Best Local Similarity 92.2%; Pred. No. 2.9e-50;
Matches 94; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 1 KATCTTHDSDPADLIEANLLWRQMGGNITRVESENKVILDSFDPTRAVEDEREISVP 60
Db 2212 KATCTTHDSDPADLIEANLLWRQMGGNITRVESENKVILDSFDPTRAVEDEREISVA 2271
QY 61 AEILRKPKFPALPIWARPDPYNPILLESWKDPDYPVPVHG 102
Db 2272 AEILRKPKFPALPIWARPDPYNPILLESWKSPDYPVPAVHG 2313
RESULT 11
AAB36932
ID AAB36932 standard; protein; 1736 AA.
XX
AC AAB36932;
XX
DT 06-AUG-2003 (revised)
DT 27-FEB-2001 (first entry)
XX
DE Hepatitis C virus tTA gene fused to polyprotein region.
XX
XX NS3-5 polyprotein; virus activity; surrogate cell.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
XX WO200066623-A2.
XX
XX 09-NOV-2000.
XX
XX 01-MAY-2000; 2000WO-CA000496.
XX
XX 04-MAY-1999; 99US-0132360P.
XX
XX (BOEH) BOEHRINGER INGELHEIM CANADA LTD.
XX
XX Pellerin C, Lamarre D;
XX
XX WPI; 2001-007207/01.
XX
XX N-PSDB; AAC83408.
XX
XX Surrogate cell-based system for assaying hepatitis C virus (HCV) NS3
FT protease activity, comprises chimeric DNA containing a transactivator
FT domain fused downstream of a HCV DNA molecule encoding NS3-5 polyprotein.
XX
XX Claim 9; Page 63-68; 70pp; English.
XX
XX The present invention relates to a surrogate cell-based system to

DR WPI; 2003-381708/36.
DR N-PSDB; ACA61114.
XX
PT New nucleic acid encoding a Met-NS3-NS4A-NS4B-NS5A-NS5B polypeptide,
PT useful as a component of an adenovector or DNA plasmid vaccine for
PT preventing or treating hepatitis C virus.
XX
XX
PS Example 1; Fig 5; 231pp; English.
XX
CC The invention relates to a nucleic acid encoding a Met-NS3-NS4A-NS4B-NS5A
CC -NS5B (NS stands for non-structural protein from HCV) polypeptide
CC appearing as AB09574. The encoded polypeptide has sufficient protease
CC activity to process itself to produce an NS5B protein that is
CC enzymatically inactive. Also included are a cultured recombinant cell
CC comprising the novel nucleic acid, making an adenovector (comprising: (a)
CC producing an adenovirus genome plasmid comprising a gene expression
CC cassette by homologous recombination between the novel nucleic acid and a
CC nucleic acid comprising a first adenovirus region from base pair 1-450
CC corresponding to either Ad5 or Ad6, a second adenovirus region from base
CC pair 3511-5548 corresponding to Ad5 or from base pair 3508-5541
CC corresponding to Ad6, joined to the first region, a third adenovirus
CC region from base pair 5549-28133 corresponding to Ad5 or from base pair
CC 5542-28156 corresponding to Ad6, joined to the second region, a fourth
CC adenovirus region from base pair 30818-33966 corresponding to Ad5 or from
CC base pair 30789-33784 corresponding to Ad6, joined to the third region,
CC and a fifth adenovirus region from base pair 33967-35935 corresponding to
CC Ad5 or from base pair 33785-35759 corresponding to Ad6, joined to the
CC fourth region; and (b) rescuing the recombinant adenovirus from the
CC recombinant adenovirus plasmid), an adenoviral vector that is produced by
CC method above appearing as ACA6113 which has a humanised version of the
CC polynucleotide of the invention and encodes the HCV inactivated
CC polypeptide, a recombinant nucleic acid comprising one or more Ad6
CC regions and a region not present in Ad6, where at least one Ad6 region is
CC selected from E1A, E1B, E2B, E2A, E4, L1, L2, L4 and L5. The nucleic acid
CC is useful as a component of an adenoviral vector or a DNA plasmid vaccine
CC providing a broad range of antigens for generating an HCV-specific cell
CC mediated immune response. The nucleic acid may also be used in treating
CC patients infected with HCV. The present sequence is the wild-type HCV
CC polypeptide with an active NS5B protein
XX
XX
SQ Sequence 1985 AA;

Query Match 92.8%; Score 514; DB 6; Length 1985;
Best Local Similarity 91.2%; Pred. No. 2.9e-50;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KATCTTHDPSDADLIEANLLWRQEMGNITRVESENKVVLDSPDPIRAVEDEREISVP 60
DB 1187 KATCTTHVSPDADLIEANLLWRQEMGNITRVESENKVVLDSPDPLRAEEDEREVSVP 1246
QY 61 AEILRPRKPPPPALPIWARPDPYNNPPLLESWKDPDYVPPVHHG 102
DB 1247 AEILRSKKFPFAAMPPIWARPDPYNNPPLLESWKDPDYVPPVHHG 1288

RESULT 14
ADR38450
ID ADR38450 standard; protein; 1985 AA.
AC ADR38450;
XX
XX 18-NOV-2004 (first entry)
XX Hepatitis C virus (isolate BK) NS3-NS5B polypeptide mutant Seq 1.
XX Hepatitis C virus; HCV; NS3; NS5A; non-structural protein;
KW virus replication activity; mutant; mutcin.
XX Hepatitis C virus (isolate BK).
OS Synthetic.
OS
FH Key Location/Qualifiers
PT Protein 2. .632

FT Misc-difference 471 /note= "NS3"
FT corresponds to poen 470 of NS3"
FT Protein 633. .686
FT Protein 687. .947
FT Protein 948. .1394
FT Misc-difference 1179 /note= "NS5A"
FT corresponds to poen 232 of NS5A"
FT Protein 1395. .1985
FT /note= "NS5B"
XX WO2004074507-A2.
XX 02-SEP-2004.
XX 09-FEB-2004; 2004WO-US003726.
XX 13-FEB-2003; 2003US-0447318P.
XX (MERI) MERCK & CO INC.
XX Grobler J, Flores O, Markel EJ;
DR WPI; 2004-635590/61.
DR N-PSDB; ADR38452.
XX Making Hepatitis C virus (HCV) replicon having increased replication
XX activity, useful in HCV research, comprises modifying HCV replicon
XX construct to encode an amino acid substitution at a position
XX corresponding to amino acid 470 of NS3.
XX
XX Claim 9; SEQ ID NO 1; 54pp; English.
XX This invention relates to a novel method for producing a Hepatitis C
XX virus (HCV) replicon having an increased replication activity.
XX Specifically, it refers to modifying an HCV replicon construct to encode
XX an amino acid substitution in NS3 (a non-structural protein that along
XX with NS4A, NS4B, NS5A and NS5B make up the virus replication machinery
XX released in the form of a polyprotein). The present invention describes
XX an amino acid substitution at a position corresponding to amino acid 470
XX of NS3 alone, or in combination with, an isoleucine in a position
XX corresponding to amino acid 232 of NS5A that confers improved cell
XX culture replication activity compared to wild type HCV. The method is
XX useful for facilitating the identification of broadly efficacious
XX compounds against different HCV isolates and facilitating HCV research,
XX where compounds that inhibit HCV replication have research and
XX therapeutic applications in identifying overall efficacy and lack of
XX unacceptable toxicity. Accordingly, they can be used to treat or inhibit
XX the onset of HCV in a patient. This polypeptide sequence is the HCV NS3-
XX NS5B polypeptide based on HCV-BK with a 471M mutation (corresponds to
XX 470M of the NS3 protein) and 1179I mutation, given in an exemplification
XX of the invention.
XX
SQ Sequence 1985 AA;

Query Match 92.8%; Score 514; DB 8; Length 1985;
Best Local Similarity 91.2%; Pred. No. 2.9e-50;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KATCTTHDPSDADLIEANLLWRQEMGNITRVESENKVVLDSPDPIRAVEDEREISVP 60
DB 1187 KATCTTHVSPDADLIEANLLWRQEMGNITRVESENKVVLDSPDPLRAEEDEREVSVP 1246
QY 61 AEILRPRKPPPPALPIWARPDPYNNPPLLESWKDPDYVPPVHHG 102
DB 1247 AEILRSKKFPFAAMPPIWARPDPYNNPPLLESWKDPDYVPPVHHG 1288

Search completed: August 12, 2005, 14:15:01
Job time : 108.012 secs

```
RESULT 15
AAW01680
ID AAW01680 standard; protein; 2201 AA.
XX
AC AAW01680;
XX
DT 17-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 19-AUG-1997 (first entry)
XX
DE HCV NS2-NS5B non-structural protein.
XX
KW Hepatitis C virus; HCV; non-structural protein; NS5B; polyprotein;
KW RNA-dependent RNA polymerase; RdRp; terminal nucleotidyl transferase;
KW TNAase; method; assay; in vitro activity; therapy; inhibitor.
XX
OS Hepatitis C virus; strain BK.
XX
PN WO9637619-A1.
XX
PD 28-NOV-1996.
XX
PF 24-MAY-1996; 96WO-IT000106.
XX
PR 25-MAY-1995; 95IT-RM000343.
XX
PA (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
PI De Francesco R, Tomei L, Behrens S;
XX
DR WPI; 1997-021225/02.
XX
PT Reproducing enzymatic activities of HCV in vitro - using sequences contg.
PT NS5B for RNA-dependent RNA polymerase and terminal nucleotidyl
PT transferase activities.
XX
PS Example 1; Page 26-34; 49pp; English.
XX
CC A novel method for reproducing in vitro the RNA-dependent RNA polymerase
CC (RdRp) activity or the terminal nucleotidyl transferase (TNAase) activity
CC encoded by hepatitis C virus (HCV), is characterised in that sequences
CC contg. NS5B are used in the mixture. The method is used for assaying in
CC vitro the activities of RdRp and TNAase encoded by HCV in order to
CC identify, for therapeutic purposes, compounds that inhibit these
CC enzymatic activities and therefore might interfere with the replication
CC of the HCV. The present sequence comprises amino acids 810-3010 of the
CC HCV polyprotein and corresponds to NS2-NS5B proteins. cDNA encoding this
CC protein was cloned between the NcoI and HindIII sites of pBlueBacIII to
CC form pBac25. Another expression plasmid, pBac5B (containing cDNA encoding
CC amino acids 2420-3010 of HCV; see AAW01679) was also constructed.
CC Extracts of Bac25- or Bac5B-infected Sf9 cells contain a novel magnesium-
CC dependent enzymatic activity that catalyses de novo RNA synthesis. This
CC activity was shown to be dependent on the presence of RNA, but
CC independent of an added primer or of the origin of the input RNA
CC molecule. As the products generated by extracts of Sf9 cells infected
CC with either Bac25 or Bac5 appeared to be identical, the experiments
CC indicated that the observed RdRp activity is encoded by the HCV NS5B
CC protein. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 17-OCT
CC -2003 to standardise OS field)
XX
SQ Sequence 2201 AA;

Query Match 92.8%; Score 514; DB 2; Length 2201;
Best Local Similarity 91.2%; Pred. No. 3.3e-50;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KATCTTHDSDPADLLEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDERISVP 60
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1403 KATCTTHVSPDADLLEANLLWRQMGNGNITRVESENKVVILDSFDPPLRAEEDEREVSVP 1462

QY 61 AEILRKPRKPPALPIWARDYNNPPLLESWKDPDYPPVPHG 102
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1463 AEILRKSKFPAAPIWARDYNNPPLLESWKDPDYPPVPHG 1504
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:04:54 ; Search time 28.57 Seconds
(without alignments)
266.510 Million cell updates/sec

Title: US-09-758-308-5
Perfect score: 554
Sequence: 1 KATCTTHDPSADLIEANL.....NPPLLESWKDPDVPVPHVHG 102

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues
Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfilea1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	521	94.0	392	2	US-08-537-811-48
2	514	92.8	997	1	US-08-324-977-50
3	514	92.8	997	2	US-08-384-616-50
4	514	92.8	997	2	US-08-904-686A-50
5	514	92.8	997	3	US-09-315-850-50
6	514	92.8	2201	3	US-08-952-981A-2
7	514	92.8	2620	1	US-08-324-977-32
8	514	92.8	2620	2	US-08-384-616-32
9	514	92.8	2620	3	US-08-904-686A-32
10	514	92.8	2620	3	US-09-315-850-32
11	514	92.8	2621	1	US-08-324-977-36
12	514	92.8	2621	2	US-08-384-616-36
13	514	92.8	2621	2	US-08-904-686A-36
14	514	92.8	2621	3	US-09-315-850-36
15	514	92.8	3010	1	US-08-324-977-2
16	514	92.8	3010	1	US-08-324-977-14
17	514	92.8	3010	2	US-08-384-616-2
18	514	92.8	3010	2	US-08-384-616-14
19	514	92.8	3010	2	US-08-904-686A-2
20	514	92.8	3010	2	US-08-904-686A-14
21	514	92.8	3010	3	US-09-315-850-2
22	514	92.8	3010	3	US-09-315-850-14
23	511	92.2	449	4	US-09-194-949A-13
24	507	91.5	1985	4	US-09-539-601-9
25	507	91.5	1985	4	US-09-539-601-12
26	507	91.5	1985	4	US-09-539-601-18
27	507	91.5	1985	4	US-09-539-601-24

28	507	91.5	2201	4	US-09-539-601-6	Sequence 6, Appli
29	507	91.5	2201	4	US-09-539-601-15	Sequence 15, Appl
30	507	91.5	3010	4	US-09-539-601-3	Sequence 3, Appli
31	507	91.5	3010	4	US-09-539-601-21	Sequence 21, Appl
32	507	91.5	3010	4	US-09-539-601-27	Sequence 27, Appl
33	506	91.3	3010	3	US-09-014-416-3	Sequence 3, Appli
34	502	90.6	2201	4	US-10-029-907-3	Sequence 3, Appli
35	500	90.3	1985	4	US-09-539-601-30	Sequence 30, Appl
36	500	90.3	3010	4	US-09-539-601-33	Sequence 33, Appl
37	478	86.3	3011	1	US-08-188-281B-1	Sequence 1, Appli
38	478	86.3	3011	1	US-08-453-552-1	Sequence 1, Appli
39	478	86.3	3011	2	US-08-710-637-1	Sequence 1, Appli
40	478	86.3	3011	5	PCT-US93-00907-1	Sequence 1, Appli
41	478	86.3	3011	5	PCT-US94-07280-1	Sequence 1, Appli
42	478	86.3	3011	5	PCT-US95-01087-1	Sequence 1, Appli
43	475	85.7	126	3	US-08-444-818-40	Sequence 40, Appl
44	475	85.7	195	5	PCT-US94-04174-17	Sequence 17, Appl
45	475	85.7	504	1	US-07-853-985A-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-537-811-48
; Sequence 48, Application US/08537811
; Patent No. 5910405
; GENERAL INFORMATION:
; APPLICANT: CHO, JOONG WYUNG
; APPLICANT: CHOI, DEOG YOUNG
; APPLICANT: KIM, CHUN HYUNG
; APPLICANT: SO, HONG SEOB
; APPLICANT: YANG, JAE YOUNG
; APPLICANT: KIM, IN SOO
; APPLICANT: KIM, JOO HO
; TITLE OF INVENTION: IMPROVED HCV DIAGNOSTIC
; TITLE OF INVENTION: AGENTS
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,811
; FILING DATE: 24-OCT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/KR94/00040
; FILING DATE: 29-APR-1994
; APPLICATION NUMBER: KR 93-7440
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Jones, Iii, Harry C
; REGISTRATION NUMBER: 20,280
; REFERENCE/DOCKET NUMBER: 8512-037-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 392 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```
; MOLECULE TYPE: protein
; FEATURE:
; OTHER INFORMATION: KHCV NS5-1.2, Fig. 3
US-08-537-811-48

Query Match          94.0%; Score 521; DB 2; Length 392;
Best Local Similarity 94.1%; Pred. No. 1.9e-55;
Matches 96; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 KATCTTHDSDPADLIEANLLWRQMGGNITRVESENKVVILDSFDPPIRAVEDEREISVP 60
    |||||
Db 110 KATCTTHDSDPADLIEANLLWRQMGGNITRVESENKVVILDSFDPPIRAVEDEREISVP 169
    |||||

QY 61 AEILRKPRKFPALPIWAPDYNPPLESWKDPDYVPPVHG 102
    |||||
Db 170 AEILRKPRKFPALPIWAPDYNPPLESWKDPDYVPPVHG 211
    |||||

RESULT 2
US-08-324-977-50
; Sequence 50, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44mb
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142

; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 997 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-324-977-50

Query Match          92.8%; Score 514; DB 1; Length 997;
Best Local Similarity 91.2%; Pred. No. 4.9e-54;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KATCTTHDSDPADLIEANLLWRQMGGNITRVESENKVVILDSFDPPIRAVEDEREISVP 60
    |||||
Db 199 KATCTTHVSPDADLIEANLLWRQMGGNITRVESENKVVILDSFDPPIRAVEDEREISVP 258
    |||||

QY 61 AEILRKPRKFPALPIWAPDYNPPLESWKDPDYVPPVHG 102
    |||||
Db 259 AEILRKPRKFPALPIWAPDYNPPLESWKDPDYVPPVHG 300
    |||||

RESULT 3
US-08-384-616-50
; Sequence 50, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
```

```
/ INFORMATION FOR SEQ ID NO: 50:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 997 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-384-616-50

Query Match          92.8%; Score 514; DB 2; Length 997;
Best Local Similarity 91.2%; Pred. No. 4.9e-54;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KATCTTHDPSDADLIEANLLWRQMGNIITRVESENKVWLDSPDPIRAVEDEREISVP 60
Db 199 KATCTTHVSPDADLIEANLLWRQMGNIITRVESENKVWLDSPDLRAEEDEREVSVP 258

QY 61 AEILRKPRKPPALPIWARPDPNPPLESWKDPDYVPPVHG 102
Db 259 AEILRSKRKPPAAMPPIWARPDPNPPLESWKDPDYVPPVHG 300

RESULT 4
US-08-904-686A-50
/ Sequence 50, Application US/08904686A
/ Patent No. 5998130
/ GENERAL INFORMATION:
/ APPLICANT: OKAYAMA, Hiroto
/ APPLICANT: FUKU, Isao
/ APPLICANT: MORI, Chisato
/ APPLICANT: TAKAMIZAWA, Akahisa
/ APPLICANT: YOSHIDA, Iwao
/ TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
/ TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
/ NUMBER OF SEQUENCES: 50
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Armstrong, Westerman, Hattori, McLeand &
/ STREET: 1725 K St. N.W. Suite 1000
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20006
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
/ SOFTWARE: ASCII
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/904,686A
/ FILING DATE: 01-AUG-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/324,977
/ FILING DATE: 18-OCT-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 2-167466
/ FILING DATE: 25-JUN-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 2-230921
/ FILING DATE: 31-AUG-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 2-305605
/ FILING DATE: 09-NOV-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/099,706
/ FILING DATE: 30-JUL-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/769,996
/ FILING DATE: 02-OCT-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/635,451
/ FILING DATE: 28-DEC-1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McLeand, Le-Nhung
```

```
/ REGISTRATION NUMBER: 31,541
/ REFERENCE/DOCKET NUMBER: 900703G
/ TELEPHONE: (202) 659-2930
/ TELEFAX: (202) 887-0357
/ INFORMATION FOR SEQ ID NO: 50:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 997 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-904-686A-50

Query Match          92.8%; Score 514; DB 2; Length 997;
Best Local Similarity 91.2%; Pred. No. 4.9e-54;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KATCTTHDPSDADLIEANLLWRQMGNIITRVESENKVWLDSPDPIRAVEDEREISVP 60
Db 199 KATCTTHVSPDADLIEANLLWRQMGNIITRVESENKVWLDSPDLRAEEDEREVSVP 258

QY 61 AEILRKPRKPPALPIWARPDPNPPLESWKDPDYVPPVHG 102
Db 259 AEILRSKRKPPAAMPPIWARPDPNPPLESWKDPDYVPPVHG 300

RESULT 5
US-09-315-850-50
/ Sequence 50, Application US/09315850
/ Patent No. 6217872
/ GENERAL INFORMATION:
/ APPLICANT: OKAYAMA, Hiroto
/ APPLICANT: FUKU, Isao
/ APPLICANT: MORI, Chisato
/ APPLICANT: TAKAMIZAWA, Akahisa
/ APPLICANT: YOSHIDA, Iwao
/ TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
/ TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
/ NUMBER OF SEQUENCES: 50
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Armstrong, Westerman, Hattori, McLeand &
/ STREET: 1725 K St. N.W. Suite 1000
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20006
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
/ SOFTWARE: ASCII
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/315,850
/ FILING DATE:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/904,686
/ FILING DATE: 01-AUG-1997
/ APPLICATION NUMBER: US 08/324,977
/ FILING DATE: 18-OCT-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 2-167466
/ FILING DATE: 25-JUN-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 2-230921
/ FILING DATE: 31-AUG-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 2-305605
/ FILING DATE: 09-NOV-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/099,706
/ FILING DATE: 30-JUL-1993
/ PRIOR APPLICATION DATA:
```

```

; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 997 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-315-850-50

Query Match          92.8%; Score 514; DB 3; Length 997;
Best Local Similarity 91.2%; Pred. No. 4.9e-54;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESKVKVILDSFDPPIRAVEDEREISVP 60
Db 199 KATCTTHVSPDADLIEANLLWRQMGNGNITRVESKVKVILDSFDPPLRAEEDEREVSVP 258

QY 61 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVPHG 102
Db 259 AEILRSKKFPAAMPPIWARPDPYNPPLLESWKDPDYVPPVPHG 300
```

RESULT 6

```

; US-08-952-981A-2
; Sequence 2, Application US/08952981A
; Patent No. 6383768
; GENERAL INFORMATION:
; APPLICANT: TOMEL, Licia
; APPLICANT: BEHRNS, Sven-Erik
; TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE RNA-DEPENDENT RNA
; TITLE OF INVENTION: POLYMERASE AND TERMINAL NUCLEOTIDYL TRANSFERASE
; FILE REFERENCE: IT0002P
; CURRENT APPLICATION NUMBER: US/08/952,981A
; CURRENT FILING DATE: 1998-03-23
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2201
; TYPE: PRT
; ORGANISM: cdna clone pCD (38-9.4)
; US-08-952-981A-2
```

```

Query Match          92.8%; Score 514; DB 3; Length 2201;
Best Local Similarity 91.2%; Pred. No. 1.4e-53;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESKVKVILDSFDPPIRAVEDEREISVP 60
Db 1403 KATCTTHVSPDADLIEANLLWRQMGNGNITRVESKVKVILDSFDPPLRAEEDEREVSVP 1462

QY 61 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVPHG 102
Db 1463 AEILRSKKFPAAMPPIWARPDPYNPPLLESWKDPDYVPPVPHG 1504
```

RESULT 7

```

; US-08-324-977-32
; Sequence 32, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
```

```

; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2620 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-324-977-32

Query Match          92.8%; Score 514; DB 1; Length 2620;
Best Local Similarity 91.2%; Pred. No. 1.8e-53;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESKVKVILDSFDPPIRAVEDEREISVP 60
Db 1823 KATCTTHVSPDADLIEANLLWRQMGNGNITRVESKVKVILDSFDPPLRAEEDEREVSVP 1882

QY 61 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVPHG 102
Db 1883 AEILRSKKFPAAMPPIWARPDPYNPPLLESWKDPDYVPPVPHG 1924

RESULT 8
US-08-384-616-32
```

; Sequence 32, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2620 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-384-616-32

Query Match 92.8%; Score 514; DB 2; Length 2620;
Best Local Similarity 91.2%; Pred. No. 1.8e-53;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
Qy 1 KATCTTHSDPADLIEANLLWRQEMGNITRVESENKVVLDSPDPIRAVEDEREISVP 60
Db 1823 KATCTTHVSPDADLIEANLLWRQEMGNITRVESENKVVLDSPDPIRAVEDEREISVP 1882
Qy 61 AEILRKPKFPALPIWARPDPYPPVPHG 102
Db 1883 AEILRKPKFPALPIWARPDPYPPVPHG 1924
RESULT 9
US-08-904-686A-32

; Sequence 32, Application US/08904686A
; Patent No. 5998130
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeLeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2620 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-904-686A-32

Query Match 92.8%; Score 514; DB 2; Length 2620;
Best Local Similarity 91.2%; Pred. No. 1.8e-53;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
Qy 1 KATCTTHSDPADLIEANLLWRQEMGNITRVESENKVVLDSPDPIRAVEDEREISVP 60
Db 1823 KATCTTHVSPDADLIEANLLWRQEMGNITRVESENKVVLDSPDPIRAVEDEREISVP 1882
Qy 61 AEILRKPKFPALPIWARPDPYPPVPHG 102

Db 1883 ABILRKSXKFPAMPPIWARPDPYNPPLLESWKDPDYVPPVVG 1924

RESULT 10
US-09-315-850-32
; Sequence 32, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,850
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686
; FILING DATE: 01-AUG-1997
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2620 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-315-850-32

Query Match 92.8%; Score 514; DB 3; Length 2620;
Best Local Similarity 91.2%; Pred. No. 1.8e-53;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KATCTTHDSDPADLIEANLLWRQMGCGNITRVESENKVVILDSFDPRAVEDEREISVP 60
Db 1823 KATCTTHVSPDADLIEANLLWRQMGCGNITRVESENKVVILDSFDPRAVEDEREISVP 1882
QY 61 AEILRKRKFPFPALPIWARPDPYNPPLLESWKDPDYVPPVVG 102
Db 1883 AEILRKSXKFPAMPPIWARPDPYNPPLLESWKDPDYVPPVVG 1924
RESULT 11
US-08-324-977-36
; Sequence 36, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2621 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-324-977-36

Query Match 92.8%; Score 514; DB 1; Length 2621;
Best Local Similarity 91.2%; Pred. NO. 1.8e-53;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KATCTTHSDPADLIEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDEREISVP 60
Db 1823 KATCTTHSVSPDADLIEANLLWRQMGNGNITRVESENKVVVLDSPDLRAEEDEREVSVP 1882

Qy 61 AEILRKPRKPPALPIWARDYNPPLESWKDPDYVPPVHVG 102
Db 1883 AEILRKSKPFAAMPPIWARDYNPPLESWKDPDYVPPVHVG 1924

RESULT 12

US-08-384-616-36
; Sequence 36, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: cDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLealand &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER: US 07/769,996
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2621 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-384-616-36

Query Match 92.8%; Score 514; DB 2; Length 2621;
Best Local Similarity 91.2%; Pred. NO. 1.8e-53;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KATCTTHSDPADLIEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDEREISVP 60
Db 1823 KATCTTHSVSPDADLIEANLLWRQMGNGNITRVESENKVVVLDSPDLRAEEDEREVSVP 1882

Qy 61 AEILRKPRKPPALPIWARDYNPPLESWKDPDYVPPVHVG 102
Db 1883 AEILRKSKPFAAMPPIWARDYNPPLESWKDPDYVPPVHVG 1924

RESULT 13

US-08-904-686A-36
; Sequence 36, Application US/08904686A
; Patent No. 598130
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: cDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLealand &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLealand, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2621 amino acids

```
;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-904-686A-36

Query Match 92.8%; Score 514; DB 2; Length 2621;
Best Local Similarity 91.2%; Pred. No. 1.8e-53;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KATCTTHDSDPADLIEANLLWRQEMGNITRVESENKVVLDSPDFPIRAVEDEREISVP 60
Db 1823 KATCTTHVSPDADLIEANLLWRQEMGNITRVESENKVVLDSPDFPLRAEEDEREVSVP 1882

Qy 61 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVHVG 102
Db 1883 AEILRKSKFPAMPPIWARPDPYNPPLLESWKDPDYVPPVHVG 1924

RESULT 14
US-09-315-850-36
; Sequence 36, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,850
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686
; FILING DATE: 01-AUG-1997
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeland, Le-Nhung
; REGISTRATION NUMBER: 31,541

;
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2621 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-315-850-36

Query Match 92.8%; Score 514; DB 3; Length 2621;
Best Local Similarity 91.2%; Pred. No. 1.8e-53;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KATCTTHDSDPADLIEANLLWRQEMGNITRVESENKVVLDSPDFPIRAVEDEREISVP 60
Db 1823 KATCTTHVSPDADLIEANLLWRQEMGNITRVESENKVVLDSPDFPLRAEEDEREVSVP 1882

Qy 61 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVHVG 102
Db 1883 AEILRKSKFPAMPPIWARPDPYNPPLLESWKDPDYVPPVHVG 1924

RESULT 15
US-08-324-977-2
; Sequence 2, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
```

```
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3010 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-324-977-2

Query Match      92.8%; Score 514; DB 1; Length 3010;
Best Local Similarity 91.2%; Pred. No. 2.2e-53;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy      1 KATCTTHDSDADLIEANLLWRQEMGNITRVESENKVVILDSFDPPIRAVEDEREISVP 60
Db      2212 KATCTTHVSPDADLIEANLLWRQEMGNITRVESENKVVILDSFDPPLRAEEDEREVSVP 2271

Qy      61 AEILRKPRKPPALPIWARDYNPPLLESWKDPDYVPPVHG 102
Db      2272 AEILRKSKKPPAAMPWARDYNPPLLESWKDPDYVPPVHG 2313
```

Search completed: August 12, 2005, 14:24:50
Job time : 29.57 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:04:14 ; Search time 20.2998 Seconds
(without alignments)
483.460 Million cell updates/sec

Title: US-09-758-308-5

Perfect score: 554

Sequence: 1 KATCTTHDPSDADLIEANL.....NPPLLESWKDPDYVPVPHVHG 102

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: Pirl:*

2: Pirl:*

3: Pirl:*

4: Pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	554	100.0	3010	1 GNVVCJ	genome polyprotein
2	518	93.5	200	2 PQ0250	polyprotein (clone
3	516	93.1	3010	1 A45573	genome polyprotein
4	514	92.8	3010	1 GNVWTC	genome polyprotein
5	511	92.2	3010	1 S18030	genome polyprotein
6	509	91.9	200	2 PQ0248	polyprotein (clone
7	506	91.3	3010	1 GNVWTV	genome polyprotein
8	475	85.7	3011	1 GNVWC3	genome polyprotein
9	472	85.2	3011	1 GNVWCH	genome polyprotein
10	471	85.0	3011	1 S40770	genome polyprotein
11	451	81.4	182	2 PQ0252	polyprotein (clone
12	409	73.8	175	2 PQ0253	polyprotein (clone
13	394	71.1	173	2 PQ0251	polyprotein (clone
14	373	67.3	3014	1 JC5620	genome polyprotein
15	336	60.6	3033	1 GNVWJ8	genome polyprotein
16	322	58.1	160	2 PQ0254	polyprotein (clone
17	318	57.4	3033	1 JQ1303	genome polyprotein
18	283	51.1	153	2 PQ0255	polyprotein (clone
19	270	48.7	103	2 S20763	nonstructural prot
20	267	48.2	89	2 S10002	CS-2 protein - hep
21	75	13.5	211	2 C84681	hypothetical prote
22	73	13.2	836	2 D87084	phenylalanyl-tRNA
23	71.5	12.9	547	2 T48551	fructosidase-like
24	71	12.8	1985	2 S19511	hypothetical prote
25	70.5	12.7	365	2 AF2893	glycosyltransferas
26	70.5	12.7	440	2 T50912	hypothetical prote
27	70.5	12.7	2599	2 A96616	unknown protein F1
28	69.5	12.5	259	2 B72505	hypothetical prote
29	69.5	12.5	490	2 A96800	hypothetical prote

30 69.5 12.5 584 2 T06163
31 69 12.5 309 2 S08343
32 69 12.5 831 2 E70620
33 68.5 12.4 374 1 BVSCES
34 68.5 12.4 466 2 G72603
35 68.5 12.4 633 2 C70358
36 67.5 12.2 340 2 C69466
37 67.5 12.2 398 2 A95870
38 67.5 12.2 450 2 D71101
39 67.5 12.2 753 2 T46614
40 67 12.1 432 2 S24332
41 67 12.1 554 2 F86244
42 67 12.1 603 2 T37518
43 67 12.1 1844 2 S01956
44 66.5 12.0 1041 2 E70760
45 66 11.9 239 2 F85774

ALIGNMENTS

RESULT 1

GNWVCJ

genome polyprotein - hepatitis C virus (strain J)

N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructur

protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

C;Species: hepatitis C virus

C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004

C;Accession: A39253; PS0086

R;Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimo

proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990

A;Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients

A;Reference number: A39253; MUID:91088550; PMID:2175903

A;Accession: A39253

A;Molecule type: Genomic RNA

A;Residues: 1-3010 <KAT>

A;Cross-references: UNIPROT:P26662; GB:D90208; NID:g221610; PIDN:BAA14233.1; PID:g22161

R;Kato, N.; Ohkoshi, S.; Shimotohno, K.

Proc. Jpn. Acad. 65B, 219-223, 1989

A;Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence var

A;Reference number: PS0085

A;Accession: PS0086

A;Molecule type: Genomic RNA

A;Residues: 2650-2707 <KA2>

A;Experimental source: Japanese isolate

C;Comment: The cleavage sites of this polyprotein have not been determined.

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; seri

F;2-115/Product: capsid protein C #status predicted <CPC>

F;116-191/Product: envelope protein M #status predicted <EPM>

F;192-389/Product: major envelope protein E #status predicted <MEE>

F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>

F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>

F;1007-1615/Product: hepatitis virus #status predicted <NS3>

F;1230-1237/Region: nucleotide-binding motif A (P-loop)

F;1312-1317/Region: nucleotide-binding motif B

F;1316-1319/Region: DEXH motif

F;1616-1862/Product: nonstructural protein NS4a #status predicted <NS4a>

F;1863-2013/Product: nonstructural protein NS4b #status predicted <NS4b>

F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

F;196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,

Query Match 100.0%; Score 554; DB 1; Length 3010;

Best Local Similarity 100.0%; Pred. No. 4.4e-48;

Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATCTTHDPSDADLIEANLWRQMGNTITRVESENKVVILDSFDFPRAVEREISVP 60

DB 2212 KATCTTHDPSDADLIEANLWRQMGNTITRVESENKVVILDSFDFPRAVEREISVP 2271

QY 61 AEILRKPRKPPALPIWARDYNNPPLLESWKDPDYVPVPHVHG 102

DB 2272 AEILRKPRKPPALPIWARDYNNPPLLESWKDPDYVPVPHVHG 2313

```
RESULT 2
PQ0250
polyprotein (clone 2206) - hepatitis C virus (isolate 22) (fragment)
C:Species: hepatitis C virus
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: PQ0250
R:Oshima, M.; Tsuchiya, M.; Yagasaki, M.; Orita, T.; Hasegawa, M.; Tomonoh, K.; Kojima,
J. Gen. Virol. 72, 2805-2809, 1991
A:Title: cDNA clones of Japanese hepatitis C virus genomes derived from a single patient
A:Reference number: PQ0245; MUID:92044457; PMID:1658209
A:Accession: PQ0250
A:Molecule type: Genomic RNA
A:Residues: 1-200 <OSH>
A:Cross-references: UNIPROT:Q9DTE3
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: polyprotein

Query Match 93.5%; Score 518; DB 2; Length 200;
Best Local Similarity 90.2%; Pred. No. 7.7e-46;
Matches 92; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESKNVILDSFDPPIRAVEDERISVP 60
DB 93 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESKNVILDSFDPPLRAEEDEREVSVA 152

QY 61 AEILRKPRKFPALPIWARPDPYNPPLLESWKDPDYVPPVVG 102
DB 153 AEILRTRRFPFAIPWARPDPYNPPLLESWKDPDYVPPVVG 194

RESULT 3
A45573
Genome polyprotein - hepatitis C virus (strain JT)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome polyprotein
Protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C>Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: A45573
R:Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata,
Virus Res. 23, 39-53, 1992
A:Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: a
A:Reference number: A45573; MUID:92295714; PMID:1318627
A:Accession: A45573
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3010 <TAN>
A:Cross-references: UNIPROT:Q00269; GB:D11168; GB:D01171; NID:G221612; PIDN:BAA01943.1;
A:Experimental source: HCV-JT
A>Note: sequence extracted from NCBI backbone (NCBI:106206, NCBIP:106207)
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <MEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polyprotein NS2 #status predicted <NS2>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: DEXH motif
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,2222

Query Match 93.1%; Score 516; DB 1; Length 3010;
Best Local Similarity 92.2%; Pred. No. 3.5e-44;
Matches 94; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESKNVILDSFDPPIRAVEDERISVP 60
DB 2212 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESKNVILDSFDPPLRAEEDEREVSVA 2271
```

```
QY 61 AEILRKPRKFPALPIWARPDPYNPPLLESWKDPDYVPPVVG 102
DB 2272 AEILRSKKFPALPIWARPDPYNPPLLESWKSPDYVPPAVHG 2313

RESULT 4
GNWVTC
Genome polyprotein - hepatitis C virus
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome polyprotein
Protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: A38465
R:Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.;
J. Virol. 65, 1105-1113, 1991
A:Title: Structure and organization of the hepatitis C virus genome isolated from human
A:Reference number: A38465; MUID:91140698; PMID:1847440
A:Accession: A38465
A:Molecule type: Genomic RNA
A:Residues: 1-3010 <TAK>
A:Cross-references: UNIPROT:P26663; EMBL:M58335; NID:G329770; PIDN:AAA72945.1; PID:G329.
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructur
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <MEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polyprotein NS2 #status predicted <NS2>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: DEXH motif
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,2222

Query Match 92.8%; Score 514; DB 1; Length 3010;
Best Local Similarity 91.2%; Pred. No. 5.6e-44;
Matches 93; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KATCTTHDSDPADLIEANLLWRQMGNGNITRVESKNVILDSFDPPIRAVEDERISVP 60
DB 2212 KATCTTHVSPDADLIEANLLWRQMGNGNITRVESKNVILDSFDPPLRAEEDEREVSVP 2271

QY 61 AEILRKPRKFPALPIWARPDPYNPPLLESWKDPDYVPPVVG 102
DB 2272 AEILRSKKFPAPPIWARPDPYNPPLLESWKDPDYVPPVVG 2313

RESULT 5
S18030
Genome polyprotein - hepatitis C virus (isolate JK1)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome polyprotein
Protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
A:Variety: isolate JK1
C>Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: S18030; S33570; A48332; S18029
R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A:Description: A whole genome of hepatitis C virus cDNA was isolated from a single pati
A:Reference number: S18028
A:Accession: S18030
A:Molecule type: Genomic RNA
A:Residues: 1-3010 <HON>
A:Cross-references: UNIPROT:Q68949; EMBL:X61596; NID:G59478; PIDN:CAA43793.1; PID:G5947
A:Experimental source: isolate JK1 from an individual
R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A:Title: Sequence analysis of putative structural regions of hepatitis C virus isolated
A:Reference number: A48332; MUID:93119270; PMID:8380322
```

A;Accession: S33570
 A;Molecule type: genomic RNA
 A;Residues: 1-547, 'T', 549-621, 'V', 623-624, 'S', 626-652, 'DL', 655-761, 'T', 763-782 <HOW>
 A;Cross-references: EMBL:X61591
 A;Note: this sequence is inconsistent with the nucleotide translation
 A;Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 320 as Trp, and TTC for residue 771 as Ser
 A;Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBI:P121748)
 C;Superfamily: hepatitis C virus genome polyprotein
 C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
 F12-115/Product: capsid protein C #status predicted <CPC>
 F116-191/Product: envelope protein M #status predicted <EPM>
 F192-389/Product: major envelope protein E #status predicted <MEE>
 F390-729/Product: nonstructural protein NS1 #status predicted <NS1>
 F730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
 F1007-1615/Product: hepatitis virus #status predicted <NS3>
 F1230-1237/Region: nucleotide-binding motif A (P-loop)
 F1312-1317/Region: nucleotide-binding motif B
 F1316-1319/Region: DEXH motif
 F1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
 F1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
 F2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
 F196,209,234,250,305,417,423,448,532,540,556,576,623,645/Banding site: carbohydrate (As)
 Query Match 92.2%; Score 511; DB 1; Length 3010;
 Best Local Similarity 92.2%; Pred. No. 1.1e-43;
 Matches 94; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 1 KATCTTHDSDADLIEANLLWRQEMGNITRVESKNVILDSFDPPIRAVEDEREISVP 60
 DB 2212 KATCTTHDSDADLIEANLLWRQEMGNITRVESKNVILDSFDPPLRAEEDEREVSVA 2271
 QY 61 AEILRKPRKPPALPIWARPDPNPPLESWKDPDYPPVPHG 102
 DB 2272 AEILRKSRKPPALPIWARPDPNPPLESWKDPDYPPVPHG 2313
 RESULT 6
 PQ0248
 polyprotein (clones 2232 and 2230) - hepatitis C virus (isolate 22) (fragment)
 C;Species: hepatitis C virus
 C;Date: 31-Mar-1992 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
 C;Accession: PQ0248; PQ0249
 R;Oshima, M.; Tsuchiya, M.; Yagasaki, M.; Orita, T.; Hasegawa, M.; Tomonoh, K.; Kojima, J. Gen. Virol. 72, 2805-2809, 1991
 A;Title: cDNA clones of Japanese hepatitis C virus genomes derived from a single patient
 A;Reference number: PQ0245; MUID:92044457; PMID:1658209
 A;Accession: PQ0248
 A;Molecule type: genomic RNA
 A;Residues: 1-200 <OSH>
 A;Cross-references: UNIPROT:Q86669; GB:S66587; NID:G238880; PIDN:AA20320.1; PID:G238880
 C;Superfamily: hepatitis C virus genome polyprotein
 C;Keywords: polyprotein

Query Match 91.9%; Score 509; DB 2; Length 200;
 Best Local Similarity 90.2%; Pred. No. 6.4e-45;
 Matches 92; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
 QY 1 KATCTTHDSDADLIEANLLWRQEMGNITRVESKNVILDSFDPPIRAVEDEREISVP 60
 DB 93 KATCTTHDSDADLIEANLLWRQEMGNITRVESKNVILDSFDPPLRAEEDEREVSVA 152
 QY 61 AEILRKPRKPPALPIWARPDPNPPLESWKDPDYPPVPHG 102
 DB 153 AEILKTRRFPALPIWARPDPNPPLESWKDPDYPPVPHG 194

RESULT 7
 GNVVTV
 genome polyprotein - hepatitis C virus (strain Taiwan)
 N;Contains: capsid protein C; envelope protein M; hepatitis virus (EC 3.4.21.98) (nonstructu
 protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
 C;Species: hepatitis C virus

A;Note: host Homo sapiens (man)
 C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
 C;Accession: A40244
 R;Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
 Virology 188, 102-113, 1992
 A;Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the
 as Trp, and TTC for residue 771 as Ser
 A;Reference number: A40244; MUID:92230206; PMID:1314449
 A;Accession: A40244
 A;Molecule type: genomic RNA
 A;Residues: 1-3010 <CHE>
 A;Cross-references: UNIPROT:P29846; GB:M84754
 C;Superfamily: hepatitis C virus genome polyprotein
 C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructur
 F116-191/Product: capsid protein C #status predicted <CPC>
 F192-389/Product: major envelope protein M #status predicted <EPM>
 F390-729/Product: nonstructural protein NS1 #status predicted <NS1>
 F730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
 F1007-1615/Product: hepatitis virus #status predicted <NS3>
 F1230-1237/Region: nucleotide-binding motif A (P-loop)
 F1312-1317/Region: nucleotide-binding motif B
 F1316-1319/Region: DEXH motif
 F1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
 F1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
 F2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
 F196,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,20
 Query Match 91.3%; Score 506; DB 1; Length 3010;
 Best Local Similarity 90.2%; Pred. No. 3.7e-43;
 Matches 92; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 QY 1 KATCTTHDSDADLIEANLLWRQEMGNITRVESKNVILDSFDPPIRAVEDEREISVP 60
 DB 2212 KAACITRTPDADLIEANLLWRQEMGNITRVESKNVILDSFDPPLRAEEDEREVSVP 2271
 QY 61 AEILRKPRKPPALPIWARPDPNPPLESWKDPDYPPVPHG 102
 DB 2272 AEILRKSRKPPALPIWARPDPNPPLESWKDPDYPPVPHG 2313
 RESULT 8
 GNVVCS
 genome polyprotein - hepatitis C virus (strain HCV-1)
 N;Contains: capsid protein C; envelope protein M; hepatitis virus (EC 3.4.21.98) (nonstruct
 protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
 C;Species: hepatitis C virus
 C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
 C;Accession: A39166; PQ0403; PQ0404
 R;Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Co
 Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
 A;Title: Genetic organization and diversity of the hepatitis C virus.
 A;Reference number: A39166; MUID:91172826; PMID:1848704
 A;Accession: A39166
 A;Molecule type: mRNA
 A;Residues: 1-3011 <CHO>
 A;Cross-references: UNIPROT:P26664; GB:M62321; NID:G329873; PIDN:AAA45676.1; PID:G32987
 R;Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.
 J. Gen. Virol. 73, 1131-1141, 1992
 A;Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to
 A;Reference number: PQ0393; MUID:92268871; PMID:1316939
 A;Accession: PQ0403
 A;Molecule type: genomic RNA
 A;Residues: 1577-1633 <CHA>
 A;Cross-references: DDBJ:D10128
 A;Experimental source: isolates E-b16
 A;Accession: PQ0404
 A;Status: preliminary
 A;Molecule type: genomic RNA
 A;Residues: 1577-1633 <CH2>
 A;Experimental source: isolates E-b17
 C;Superfamily: hepatitis C virus genome polyprotein
 C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructur
 F1-115/Product: capsid protein C #status predicted <CPC>


```
Query Match      71.1%; Score 394; DB 2; Length 173;  
Best Local Similarity 88.9%; Pred.No.3.5e-33;  
Matches 72; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
```

Qy 1 KATCTTHDSPADLIEANLLWRQMCGNITRVSENNKVILDSFDPIRAVEDREISVP 60
 ||||| :||||| :||||| :||||| :||||| :||||| :
Db 93 KATCTTHDSPDADLIEANLLWRQMCGNITRVSENNKVILDSFDPLRABEDEREVSVA 152

Qy 61 AEILRKPKFPFPAIPWARPD 81
 ||||| :||||| :||||| :||||| :
Db 153 AEILKRRTRFRFPFAIPVWARPD 173

Qy 61 ABLRKRKPPALPIIWARDYNNPPLLESWKDDPYDPVHVHG 102
 : : : : : : : : : : :
Db 2273 ADCFKGPAFPALPVWARPGYDPPILLETWKRDPYPQVWG 2314

RESULT 15
GNWJ78
genome polyprotein - hepatitis C virus (strain HC-J8)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: A40250; PQ0397; PQ0559
R:Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fukuda, S
Virology 188, 331-341, 1992
A:Title: Full-length sequence of a hepatitis C virus genome having poor homology to ref
A:Reference number: A40250; MUID:92230232; PMID:1314459
A:Accession: A40250
A:Molecule type: genomic RNA
A:Residues: 1-3033 <OK>
A:Cross-references: UNIPROT:P26661; GB:D10988; GB:D01221; NID:G221608; PIDN:BAA01761.1;
R:Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Feuchterer, J.F.; Follett, E.; Yap, P.
J. Gen. Virol. 73, 1131-1141, 1992
A:Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to
A:Reference number: PQ0393; MUID:92268871; PMID:1316939
A:Accession: PQ0397
A:Molecule type: genomic RNA
A:Residues: 2678-2754 <CHA>

A;Cross-references: DBJ:DI0134
A;Experimental source: isolate E-b12
R;Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno
Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A;Title: Distribution of plural HCV types in Japan.
A;Reference number: PQ0554; MUID:92068204; PMID:1720309
A;Accession: PQ0559
A;Molecule type: mRNA
A;Residues: 2678-2729 <KAT>
A;Cross-references: GB:DI0562; GB:D90518; NID:G21523; PIDN:BAA01418.1; PID:G221524
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructure
F;1-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEES>
F;390-733/Product: nonstructural protein NS1 #status predicted <NS1>
F;734-1010/Product: nonstructural protein NS2 #status predicted <NS2>
F;1011-1619/Product: hepatitis virus #status predicted <NS3>
F;1234-1241/Region: nucleotide-binding motif A (P-loop)
F;1316-1321/Region: nucleotide-binding motif B
F;1320-1323/Region: DEXH motif
F;1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
F;2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,233,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038,23

Query Match 60.6%; Score 336; DB 1; Length 3033;
Best Local Similarity 60.8%; Pred. No. 1.1e-25;
Matches 62; Conservative 16; Mismatches 20; Indels 4; Gaps 1;
QY 1 KATCTTHDSDPADLIEANLLMRQMGNIITVSESKVIVLDSFDPPIRAVEDEREISVP 60
DB 2216 KATCTTHKTA YDCDWDALF---MGDVTIESDSKVIVLDSLDSMTVEDDREPSVP 2271
QY 61 AEILRKPRKPPALPIWARPDPNPPLESWKOPDYVPPVHG 102
DB 2272 SEYLIKRRKFPFPPALPPWARPDPNPNVLIETWKRPGYEPPTVLG 2313

Search completed: August 12, 2005, 14:22:52
Job time : 22.2998 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2005, 14:03:18 ; Search time 94.231 Seconds
(without alignments)
554.298 Million cell updates/sec

Title: US-09-758-308-5
Perfect score: 554
Sequence: 1 KATCTTHSPDADLIEANL.....NPPLLESWKDPDYVPVHHG 102

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	554	100.0	3010	1 P09JHCVJA	P26662 h genome po
2	545	98.6	3010	2 Q9J310	Q9J310 hepatitis c
3	545	98.4	447	2 Q8B0C7	Q8B0C7 hepatitis c
4	536	96.8	254	2 Q6TZB5	Q6TZB5 hepatitis c
5	536	96.8	254	2 Q6TZK2	Q6TZK2 hepatitis c
6	536	96.8	254	2 Q6TZV7	Q6TZV7 hepatitis c
7	535	96.6	447	2 Q89489	Q89489 hepatitis c
8	535	96.6	3008	2 Q9J3F4	Q9J3F4 hepatitis c
9	534	96.4	254	2 Q6TZL9	Q6TZL9 hepatitis c
10	533	96.2	447	2 Q8B0G3	Q8B0G3 hepatitis c
11	532	96.0	254	2 Q6TZD4	Q6TZD4 hepatitis c
12	532	96.0	254	2 Q6TZD6	Q6TZD6 hepatitis c
13	532	96.0	254	2 Q6TZD7	Q6TZD7 hepatitis c
14	532	96.0	254	2 Q6TZD2	Q6TZD2 hepatitis c
15	532	96.0	254	2 Q6TZB3	Q6TZB3 hepatitis c
16	532	96.0	254	2 Q6TZF2	Q6TZF2 hepatitis c
17	532	96.0	254	2 Q6TZF6	Q6TZF6 hepatitis c
18	532	96.0	254	2 Q6TZG0	Q6TZG0 hepatitis c
19	532	96.0	254	2 Q6TZG4	Q6TZG4 hepatitis c
20	532	96.0	254	2 Q6TZG6	Q6TZG6 hepatitis c
21	532	96.0	254	2 Q6TZG7	Q6TZG7 hepatitis c
22	532	96.0	254	2 Q6TZG8	Q6TZG8 hepatitis c
23	532	96.0	254	2 Q6TZG9	Q6TZG9 hepatitis c
24	532	96.0	254	2 Q6TZH0	Q6TZH0 hepatitis c
25	532	96.0	254	2 Q6TZH1	Q6TZH1 hepatitis c
26	532	96.0	254	2 Q6TZH2	Q6TZH2 hepatitis c
27	532	96.0	254	2 Q6TZH3	Q6TZH3 hepatitis c
28	532	96.0	254	2 Q6TZH4	Q6TZH4 hepatitis c
29	532	96.0	254	2 Q6TZH5	Q6TZH5 hepatitis c
30	532	96.0	254	2 Q6TZH6	Q6TZH6 hepatitis c
31	532	96.0	254	2 Q6TZH6	Q6TZH6 hepatitis c

32 532 96.0 254 2 Q6TZH8 hepatitis c
33 532 96.0 254 2 Q6TZH9 hepatitis c
34 532 96.0 254 2 Q6TZI0 hepatitis c
35 532 96.0 254 2 Q6TZI2 hepatitis c
36 532 96.0 254 2 Q6TZI7 hepatitis c
37 532 96.0 254 2 Q6TZI8 hepatitis c
38 532 96.0 254 2 Q6TZI9 hepatitis c
39 532 96.0 254 2 Q6TZJ3 hepatitis c
40 532 96.0 254 2 Q6TZJ5 hepatitis c
41 532 96.0 254 2 Q6TZJ6 hepatitis c
42 532 96.0 254 2 Q6TZJ7 hepatitis c
43 532 96.0 254 2 Q6TZJ8 hepatitis c
44 532 96.0 254 2 Q6TZK2 hepatitis c
45 532 96.0 254 2 Q6TZK3 hepatitis c

ALIGNMENTS

RESULT 1
POLG_HCVJA STANDARD; PRT; 3010 AA.
AC P26662;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
DE Envelope glycoprotein E1 (GP12) (GP15); Envelope glycoprotein E2
DE (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
DE (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
DE NS4B (P27); Nonstructural protein NS4A (P4); Nonstructural protein
DE NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
OS Hepatitis C virus (isolate Japanese) (HCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCBI_TaxID=11116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91088550; PubMed=2175903;
RA Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
RA Sugimura T., Shimotohno K.;
RT "Molecular cloning of the human hepatitis C virus genome from Japanese
patients with non-A, non-B hepatitis."
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
[2]

DISCUSSION OF SEQUENCE.

MEDLINE=91192160; PubMed=1849488; DOI=10.1016/0014-5793(91)80322-T;
Kato N., Hijikata M., Nakagawa M., Ootsuyama Y., Muraishi K.,
Ohkoshi S., Shimotohno K.;
RT "Molecular structure of the Japanese hepatitis C viral genome.";
FEBS Lett. 280:325-328(1991).
CC - FUNCTION: The small proteins NS2A, NS2B, NS4A and NS4B are
CC hydrophobic, suggesting a possible membrane-related function. NS3
CC and NS5 may play a role in the viral RNA replication.
CC - CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
CC precursor polyprotein, commonly with Asp or Glu in the P6
CC position, Cys or Thr in P1 and Ser or Ala in P1'.
CC - CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC - SUBUNIT: The virion of this virus is a nucleocapsid covered by a
CC lipoprotein envelope. The envelope consists of two proteins:
CC protein M and glycoprotein E. The nucleocapsid is a complex of
CC protein C and mRNA.
CC - SIMILARITY: Contains 1 peptidase S29 domain.
CC - SIMILARITY: Contains 1 peptidase U39 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch).

or send an email to license@ieib-sib.ch.

EMBL; D90208; BAA14233.1; -.
 HSR; A39253; GNWVCJ.
 PIR; P26663; LJXP.
 MEROPS; S29.001; -.
 MEROPS; U39.001; -.
 InterPro; IPR001410; DEAD.
 InterPro; IPR002522; HCV capsid.
 InterPro; IPR002521; HCV core.
 InterPro; IPR002519; HCV env.
 InterPro; IPR002531; HCV NS1.
 InterPro; IPR000745; HCV NS4a.
 InterPro; IPR001490; HCV NS4b.
 InterPro; IPR002868; HCV NS5a.
 InterPro; IPR002166; HCV RdRP.
 InterPro; IPR001650; Helicase C.
 InterPro; IPR009003; Pept_Ser_Cys.
 InterPro; IPR002518; Pept_U39_HCV_NS2.
 InterPro; IPR004109; Peptidase S29.
 InterPro; IPR007095; RNA_pol_DS_PS.
 InterPro; IPR007094; RNA_pol_PSVir.
 Pfam; PF01543; HCV_capsid; 1.
 Pfam; PF01542; HCV_core; 1.
 Pfam; PF01539; HCV_env; 1.
 Pfam; PF01560; HCV_NS1; 1.
 Pfam; PF01538; HCV_NS2; 1.
 Pfam; PF02907; HCV_NS3; 1.
 Pfam; PF01006; HCV_NS4a; 1.
 Pfam; PF01001; HCV_NS4b; 1.
 Pfam; PF01506; HCV_NS5a; 1.
 Pfam; PF00271; Helicase C; 1.
 Pfam; PF00998; Viral_RdRP; 1.
 SMART; SM00487; DEXDC; 1.
 ATP-binding; Coat protein; Core protein; Envelope protein; Glycoprotein; Helicase; Hydrolyase; Nonstructural protein; RNA-directed RNA polymerase; Serine protease; Transferrase; Transmembrane.
 INIT_MET 1 1 Removed from capsid protein C by the cellular aminopeptidase.
 CHAIN 1 115 Capsid protein C (Potential).
 CHAIN 116 191 Matrix protein (Potential).
 CHAIN 192 383 Major envelope protein E (Potential).
 CHAIN 384 729 Nonstructural protein NS1 (Potential).
 CHAIN 730 1006 Nonstructural protein NS2 (Potential).
 CHAIN 1007 1615 Protease/helicase NS3 (Potential).
 CHAIN 1616 1862 Nonstructural protein NS4A (Potential).
 CHAIN 1863 2013 Nonstructural protein NS4B (Potential).
 CHAIN 2014 3010 RNA-directed RNA polymerase (Potential).
 TRANSMEM 347 369 Potential.
 ACT_SITE 1083 1083 Charge relay system (By similarity).
 ACT_SITE 1107 1107 Charge relay system (By similarity).
 ACT_SITE 1165 1165 Charge relay system (By similarity).
 NP_BIND 1230 1237 ATP (Potential).
 SITE 1316 1319 DECH box.
 CARBOHYD 196 196 N-linked (GlcNAc...) (Potential).
 CARBOHYD 209 209 N-linked (GlcNAc...) (Potential).
 CARBOHYD 234 234 N-linked (GlcNAc...) (Potential).
 CARBOHYD 250 250 N-linked (GlcNAc...) (Potential).
 CARBOHYD 305 305 N-linked (GlcNAc...) (Potential).
 CARBOHYD 417 417 N-linked (GlcNAc...) (Potential).
 CARBOHYD 423 423 N-linked (GlcNAc...) (Potential).
 CARBOHYD 430 430 N-linked (GlcNAc...) (Potential).
 CARBOHYD 448 448 N-linked (GlcNAc...) (Potential).
 CARBOHYD 532 532 N-linked (GlcNAc...) (Potential).
 CARBOHYD 556 556 N-linked (GlcNAc...) (Potential).
 CARBOHYD 576 576 N-linked (GlcNAc...) (Potential).
 CARBOHYD 623 623 N-linked (GlcNAc...) (Potential).
 CARBOHYD 645 645 N-linked (GlcNAc...) (Potential).
 CARBOHYD 2041 2041 N-linked (GlcNAc...) (Potential).
 CARBOHYD 2077 2077 N-linked (GlcNAc...) (Potential).
 CARBOHYD 2240 2240 N-linked (GlcNAc...) (Potential).
 CARBOHYD 2788 2788 N-linked (GlcNAc...) (Potential).

DR Pfam; PF01006; HCV NS4a; 1.
 DR Pfam; PF01001; HCV NS4b; 1.
 DR Pfam; PF01506; HCV NS5a; 1.
 DR Pfam; PF00271; Helicase C; 1.
 DR Pfam; PF00998; Viral RdRp; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 SQ SEQUENCE 3010 AA; 326693 MW; 074098DB305AF1A9 CRC64;
 Query Match 98.4%; Score 546; DB 2; Length 3010;
 Best Local Similarity 98.0%; Pred. No. 1.8e-44;
 Matches 100; Conservative 1; Mismatches 0; Gaps 0;
 Indels 0;
 QY 1 KATCTTHDSPADLIEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDEREISVP 60
 Db 2212 KATCTTHDSPADLIEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDEREISVP 2271
 QY 61 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVHG 102
 Db 2272 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVHG 2313

RESULT 3

Q8B0C7 PRELIMINARY; PRT; 447 AA.
 AC Q8B0C7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22255437; PubMed=12368350;
 RX DOI=10.1128/JVI.76.21.11079-11090.2002;
 RA Sarrazin C., Herrmann E., Bruch K., Zeuzem S.;
 RT "Hepatitis C virus non-structural (NS)5A protein and interferon
 resistance: a new model for testing the reliability of mutational
 analyses";
 RL J. Virol. 76:11079-11090(2002).
 DR EMBL; AJ507192; CAD45235.1; -.
 DR InterPro; IPR002868; HCV_NS5a.
 DR Pfam; PF01506; HCV_NS5a; 1.
 KW Polyprotein.
 FT NON_TER 1 1
 FT CHAIN <1 >447 non-structural protein 5A.
 FT NON_TER 447 447
 SQ SEQUENCE 447 AA; 48708 MW; 8D9C204FCECB9767 CRC64;

Query Match 98.4%; Score 545; DB 2; Length 447;
 Best Local Similarity 97.1%; Pred. No. 2.7e-45;
 Matches 99; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Indels 0;
 QY 1 KATCTTHDSPADLIEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDEREISVP 60
 Db 240 KATCTTHDSPADLIEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDEREISVP 299
 QY 61 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVHG 102
 Db 300 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVHG 341

RESULT 4

Q6TZB5 PRELIMINARY; PRT; 254 AA.
 AC Q6TZB5;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE NS5A (Fragment).
 GN Name=NS5A;
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Puig-Basagoiti F., Saiz J.C., Forns X.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY381888; AAQ86478.1; -.
 DR InterPro; IPR002868; HCV_NS5a.
 DR Pfam; PF01506; HCV_NS5a; 1.
 FT NON_TER 1 1
 FT NON_TER 254 254
 SQ SEQUENCE 254 AA; 27242 MW; 2B2F15FE832154ED CRC64;
 Query Match 96.8%; Score 536; DB 2; Length 254;
 Best Local Similarity 95.1%; Pred. No. 1.1e-44;
 Matches 97; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 Indels 0;

QY 1 KATCTTHDSPADLIEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDEREISVP 60
 Db 64 KATCTTHDSPADLIEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDEREISVP 123

QY 61 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVHG 102
 Db 124 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVHG 165

RESULT 5

Q6TZX2 PRELIMINARY; PRT; 254 AA.
 AC Q6TZX2;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE NS5A (Fragment).
 GN Name=NS5A;
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Puig-Basagoiti F., Saiz J.C., Forns X.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY381881; AAQ86271.1; -.
 DR InterPro; IPR002868; HCV_NS5a.
 DR Pfam; PF01506; HCV_NS5a; 1.
 FT NON_TER 1 1
 FT NON_TER 254 254
 SQ SEQUENCE 254 AA; 27148 MW; 6D26E3F416CA8189 CRC64;

Query Match 96.8%; Score 536; DB 2; Length 254;
 Best Local Similarity 96.1%; Pred. No. 1.1e-44;
 Matches 98; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Indels 0;

QY 1 KATCTTHDSPADLIEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDEREISVP 60
 Db 64 KATCTTHDSPADLIEANLLWRQMGNGNITRVESENKVVILDSFDPPIRAVEDEREISVP 123

QY 61 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVHG 102
 Db 124 AEILRKPRKPPALPIWARPDPYNPPLLESWKDPDYVPPVHG 165

RESULT 6

Q6TZY7 PRELIMINARY; PRT; 254 AA.
 AC Q6TZY7;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)

```
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE NS5A (Fragment).
GN Name=NS5A;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Puig-Basagotti F., Saiz J.C., Forns X.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY381666; AAQ8256.1; -.
DR InterPro; IPR002868; HCV NS5a.
DR Pfam; PF01506; HCV_NS5a; 1.
FT NON_TER 1
FT NON_TER 254
SQ SEQUENCE 254 AA; 27178 MW; 31CBE3F416CA97CB CRC64;

Query Match          96.6%; Score 536; DB 2; Length 254;
Best Local Similarity 96.1%; Pred. NO. 1.1e-44;
Matches 98; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KATCTTHDPSADLLEALLNRQEMGNITRVESENKVVILDSFDPPIRAVEDEREISVP 60
Db 64 KATCTTHDPSADLLEALLNRQEMGNITRVESENKVVILDSFDPPLRAEEDEREISVP 123

QY 61 AEILRKPRFPALPIWARPDPYNPPLLESWKDPDYVPPVHHG 102
Db 124 AEILRKPRFPALPIWARPDPYNPPLLESWKDPDYVPPVHHG 165

RESULT 7
O89489 PRELIMINARY; PRT; 447 AA.
ID O89489
AC O89489;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=genotype 1b;
RX MEDLINE=98295834; PubMed=9634077;
RA Duverlie G., Khorsi H., Castelain S., Jaillon O., Izopet J., Lunel F.,
RA Eb F., Penin F., Wychowski C.;
RT "Sequence analysis of the NS5A protein of European hepatitis C virus
RT 1b isolates and relation to interferon sensitivity.";
RL J. Gen. Virol. 79:1373-1381 (1998).
DR EMBL; AF033367; AAC41009.1; -.
DR PIR; PQ0252; PQ0252.
DR InterPro; IPR002868; HCV NS5a.
DR Pfam; PF01506; HCV_NS5a; 1.
KW Polyprotein.
FT NON_TER 1
FT CHAIN <1> >447 NS5A.
FT NON_TER 447
SQ SEQUENCE 447 AA; 48946 MW; DE15F2F3C19E1B0C CRC64;

Query Match          96.6%; Score 535; DB 2; Length 447;
Best Local Similarity 96.1%; Pred. NO. 2.6e-44;
Matches 98; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KATCTTHDPSADLLEALLNRQEMGNITRVESENKVVILDSFDPPIRAVEDEREISVP 60
Db 240 KATCTTHDPSADLLEALLNRQEMGNITRVESENKVVILDSFDPPLRAEEDEREISVP 299

QY 61 AEILRKPRFPALPIWARPDPYNPPLLESWKDPDYVPPVHHG 102
```

```
Db 300 AEILRKPRFPALPIWARPDPYNPPLLESWKDPDYVPPVHHG 341

RESULT 8
Q9U3F4 PRELIMINARY; PRT; 3008 AA.
ID Q9U3F4
AC Q9U3F4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Name=MD34;
GN Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MD34;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF208024; AAF61205.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; O8JVS1; 1CWK.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0002236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; Cytc_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral_RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane
SQ SEQUENCE 3008 AA; 326835 MW; 99AE09E14C3109F4 CRC64;
```

DR	InterPro; IPR002868; HCV_NS5a.
DR	Pfam; PF01506; HCV_NS5a; 1.
KW	Polyprotein.
FT	NON_TER 1 non-structural protein 5A.
FT	CHAIN <1> >447
FT	NON_TER 447 447
FT	SEQUENCE 447 AA; 48586 MW; EE878DAEBEC97243 CRC64;
SQ	
	Query Match 96.2%; Score 533; DB 2; Length 447;
	Best Local Similarity 95.1%; Pred. No. 4.le-44;
	Matches 97; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy	1 KATCTTHSDPADLIEANLLWRQEGGNITRVESKNKVILDSFDPIRAVEDEREISVP 60 : : : : : : : :
Dd	240 KATCTTHSDPADLIEANLLWRQEGGNITRVESKNKVILDSFDPLRABEDEREVSVP 299 : : : : : : : :
Qy	61 AEILRKPKFPALPWARPDYNNPPLESWKDPDYVPPVVHG 102 : : : : : : : :
Dd	300 AEILRSKRKFPALPWIARPDYNNPPLESWKDIDYAPPVHVG 341 : : : : : : : :
RESULT 11	
Q6TZD4	PRELIMINARY; PRT; 254 AA.
ID Q6TZD4 AC Q6TZD4	
DT 05-JUL-2004 (TrEMBLrel. 27, Created)	
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	
DE NS5A (Fragment).	
DN Name=NS5A;	
OS Hepatitis C virus.	
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;	
CC Hepacivirus.	
OX NCBI_TaxId=11103;	
RN [1]	
RP SEQUENCE FROM N.A.	
RA Puig-Basagoti F., Saiz J.C., Fornes X.;	
RL Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.	
DR EMBL; AY381869; AAO86459.1; -.	
DR InterPro; IPR002868; HCV_NS5a.	
DR Pfam; PF01506; HCV_NS5a;-1.	
FT NON_TER 1 1	
FT NON_TER 254 254	
FT SEQUENCE 254 AA; 27153 MW; BABF5A73A84F5BC0 CRC64;	
SQ	
	Query Match 96.0%; Score 532; DB 2; Length 254;
	Best Local Similarity 94.1%; Pred. No. 2.8e-44;
	Matches 96; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Qy	1 KATCTTHSDPADLIEANLLWRQEGGNITRVESKNKVILDSFDPIRAVEDEREISVP 60 : : : : : : : :
Dd	64 KATCTTHSDPADLIEANLLWRQEGGNITRVESKNKVILDSFDPLRABEDEREVSVP 123 : : : : : : : :
Qy	61 AEILRKPKFPALPWARPDYNNPPLESWKDPDYVPPVVHG 102 : : : : : : : :
Dd	124 AEILRRTRKFPPAMPWIARPDYNNPPLESWKDPDYVPPVVHG 165 : : : : : : : :
RESULT 12	
Q6TZD6	PRELIMINARY; PRT; 254 AA.
ID Q6TZD6 AC Q6TZD6	
DT 05-JUL-2004 (TrEMBLrel. 27, Created)	
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	
DE NS5A (Fragment).	
DN Name=NS5A;	
OS Hepatitis C virus.	
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;	
CC Hepacivirus.	
OX NCBI_TaxId=11103;	
RN [1]	
RP SEQUENCE FROM N.A.	

[illegible]